

Constructive student feedback: Online vs. traditional course evaluations

Judy Donovan, Ed.D.
Indiana University Northwest

Cynthia E. Mader, Ph.D.
Grand Valley State University

John Shinsky, Ph.D.
Grand Valley State University

Abstract

Substantial efforts have been made recently to compare the effectiveness of traditional course formats to alternative formats (most often, online delivery compared to traditional on-site delivery). This study examines, not the delivery format but rather the evaluation format. It compares traditional paper and pencil methods for course evaluation with electronic methods. Eleven instructors took part in the study. Each instructor taught two sections of the same course; at the end, one course received an online course evaluation, the other a traditional pencil and paper evaluation. Enrollment in these 22 sections was 519 students. Researchers analyzed open-ended comments as well as quantitative rankings for the course evaluations. Researchers found no significant differences in numerical rankings between the two evaluation formats. However, differences were found in number and length of comments, the ratio of positive to negative comments, and the ratio of formative to summative comments. Students completing faculty evaluations online wrote more comments, and the comments were more often formative (defined as a comment that gave specific reasons for judgment so that the instructor knew what the student was suggesting be kept or changed) in nature.

Introduction

Faculty course evaluations have the potential to affect professional and career advancement, promotion, and tenure. Few areas in higher education cause more anxiety than course evaluations, and few areas have been studied more for validity and reliability (Wachtel, 1998). At a time when online course methods such as teaching, testing, grading, and discussion are no longer a novelty, online course evaluations bring the advantage of saving time and resources over the traditional paper and pencil scan sheet method. Thus, instructors may be encouraged or even required to use them in place of the more cumbersome paper method.

Regardless of convenience or efficiency, however, instructor resistance to online evaluations has been well documented, especially perceptions of lower return rates and higher percentage of negative responses. The first purpose of this study was to verify the results of existing research in those two areas. The second purpose was to expand on existing research in the area of open-ended responses. Despite the attention paid to them by faculty and their peers, open-ended student comments have not received the research attention we sought. These, then, were our objectives:

1. To verify research on return rate and quantitative rankings
2. To investigate open-ended responses, based on an analysis of these factors:
 - a. Extent of responses (number, length, proportion of respondents)
 - b. Nature of responses (positive or negative)
 - c. Usefulness of responses in improving instruction (summative or formative)

The second objective, investigating open-ended student comments, was the one that most intrigued the researchers and may offer the most significant and interesting findings.

Perspectives from the Literature

Previous research indicates both benefits and limitations inherent in online course evaluations. The benefits include time and cost savings, faster reporting of results, and possible improved quantity and quality of student comments (Kuhtman, 2004). In addition, online evaluations are less subject to faculty influence, allow students to take as much time as they wish to complete the evaluation, and also allow students to choose the time they wish to complete the evaluation (Anderson, Cain, & Bird, 2005).

One study reported that students preferred completing electronic instructor evaluations to paper ones (Layne, DeCristoforo, & McGinty, 1999). A focus group in another study reported that the online tool was easy to use, students liked the anonymity of the online evaluation, and the online evaluation allowed them to offer more thoughtful remarks than did the traditional, in-class, print-based teaching evaluation (Ravelli, 2002). In another study, over 90% of students marked Agree or Strongly Agree when asked if they preferred online to traditional evaluation format (Anderson et al., 2005). On the other hand, online evaluations may have disadvantages, such as requiring students to have computer access (Anderson et al., 2005).

Also documented is faculty resistance when moving from traditional to online evaluations (Sorenson & Reiner in Sorenson & Johnson, 2003). Some research found that faculty prefer traditional evaluations because they believe traditional methods produce a higher rate of return and more accurate responses (Dommeyer, Baum, Chapman & Hanna, 2002). Some faculty who are not proficient with computers or knowledgeable about online surveys also believe that online evaluations are less accurate than traditional ones (Anderson et al., 2005). Other concerns voiced by faculty who object to online evaluations include the beliefs that quantitative scores are lower, negative comments are more frequent, student return rate is lower, and while students voice more dissatisfaction with less favored instructors, they are not as motivated to express satisfaction with more favored instructors (Ravelli, 2000).

- *Return Rate.* Some studies verify concerns about return rates. Lower response rate has been linked with online evaluations (Dommeyer, et al., 2004; Layne et al., 1999; Sorenson & Johnson, 2003). Higher response rates may be helped by the use of course management systems such as Blackboard, which have been found to increase response rates (Oliver & Sautter, 2005).
- *Quantitative Rankings.* Little difference has been reported in quantitative scores. Online evaluations produce essentially the same quantitative scores as traditionally delivered evaluations (Dommeyer et al., 2004; Layne et al., 1999; Sorenson & Johnson, 2003). A recent work concluded that although online response rates were lower, the mean scores were the same as those derived from traditionally delivered evaluations (Avery, Bryant, Mathios, Kang, & Bell, 2006). It also found that smaller classes tended to have higher survey response rates, and that online survey response rates increased over time. Johnson (2002) noted that response rates increased yearly from 40%, then 51%, 62% and finally 71% the

final year of a study. Carini, Hayek, Kuh, Kennedy and Ouimet (2003), however, found that students who completed the web-based format of a survey responded more favorably toward faculty than students completing the paper format.

- *Extent of Open-Ended Responses.* Online open-ended responses have been studied for potential differences in length compared to traditional evaluation responses. Researchers report that student open-ended responses are lengthier in the online evaluation format, and more comments are made. Kasiar, Schroeder, & Holstaad (2001) reported, “The total number of words typed per student using the online system was more than 7 times that of the student using the traditional system.” Another study of a graduate management class found that students wrote an average of four times as many words online (62:15 words/student) as they did using paper and pencil evaluation formats (Hmielecki & Champagne, 2000, p. 5).
- *Nature and Usefulness of Open-Ended Responses.* The researchers found nothing that analyzed the content and substance of open-ended responses. Because these are the very responses that often draw the most attention from faculty and others during the evaluation process, this study hoped to bridge that apparent gap in the literature.

Methods and Techniques

The study was conducted at a large Midwestern public university and included 22 sections of graduate and undergraduate Education courses. A total of 30 instructors were identified who fit the study’s primary criteria: teaching two sections of the same course during the same semester. Of these 30 instructors, 11 were identified who fit the criteria and were willing to administer the online evaluation to one section and the traditional paper evaluation to the other section. Those who were unwilling offered reasons such as the following: first time teaching a course and unsure about varying evaluation format; belief that they would receive more negative evaluations online; unsure of how to administer online evaluations; and belief that return rate would be lower.

At the end of the course, one section of the instructor’s classes filled out the course evaluation in class in the traditional manner, while the instructor’s other section of the same course completed the evaluation online in electronic format. Items and questions were identical in each format. The department’s office assistants then paired and coded the completed evaluations, removed identifying course and instructor information, and provided photocopied sets to the researchers.

First, the researchers analyzed return rates and quantitative rankings. Second, they analyzed open-ended comments to determine length, number, and proportion of respondents making such comments. Finally, working in pairs, the researchers analyzed each open-ended comment and categorized it as a) positive or negative and b) formative or summative. A formative comment was defined as one that gave specific reasons for judgment so that the instructor knew what the student was suggesting be kept or changed. A summative comment was defined as one lacking detail on what to keep or change. These examples may provide clarity as to the judgments made by the researchers.

Table 1. Examples of Researcher Interpretations

Open-Ended Comment	Positive	Negative	Formative	Summative
This was a wonderful class.	x			x
Great professor!	x			x
Good job!	x			x
I appreciated how you respected everyone's opinion.	x		x	
We always had very good discussions, which made me want to come back.	x		x	
Made me feel comfortable participating.	x		x	
This class was frustrating for me at times.		x		x
This class was horrible.		x		x
This is one of the worst classes I've ever taken.		x		x
It would help if you explained assignments better.		x	x	
He needs work with computers, really struggles.		x	x	
Assignments were not clearly explained.		x	x	

No attempt was made to correlate the possible combinations of positive/negative and formative/summative. Although negative comments may be less pleasant to hear than positive comments, they can be just as instructive if they are formative in content rather than simply summative.

In some cases, student comments reflect circumstances that are beyond the instructor's control. Prior to tabulating data, the researchers determined that such comments would not be included in the data analysis. An example would be: "Classrooms in this building are always too cold." Although such comments may be useful for other reasons, they were not included in the data analysis because they did not directly pertain to instructor or course evaluation.

Researchers also had to determine how to handle multiple judgments from each student. The first printing of comments came to the researchers grouped by each student regardless of the number of sentences or judgments within the block. The researchers reformatted each set of comments into discrete judgments. For example, the following two-sentence block was determined to contain three separate judgments:

"I really enjoyed her knowledge (1). She could stand to be a little more organized, (2) but overall she was very good (3)."

Results

Return Rates

Table 2 shows the difference in return rates between the two groups. The number of total evaluations returned was 413 out of 519 students, an overall return rate of about 80%. About half, or 48% of the students who returned evaluations completed online ones, and 52% completed traditional evaluations. Of the students who filled out traditional evaluations, 83% (215 students of 258 offered traditional evaluations) returned the forms, compared with 76% (198 of 261) who submitted online evaluations. In some cases, the online evaluations generated a higher rate of return, but overall more students returned traditional evaluations than online evaluations.

Table 2. Summary Data on Return Rates (N=519)

Students per Evaluation Format (n=519)		Evaluations Returned (n=413)		% Return Rate based on Evaluations Returned		
Online	Traditional	Online	Traditional	Online	Traditional	Difference
261	258	198	215	48%	52%	4 % higher from traditional

(Note: Tables at the end of this paper contain complete data from each course section.)

Quantitative Rankings: Little difference was found in quantitative results between traditional and online evaluations. Using a five-point Likert scale, students responded to 14 questions asking the extent to which the instructor was prepared, fair, enthusiastic, etc. The responses for each of the 14 questions was averaged for each individual instructor, resulting in one average score per instructor for each question. When responses were tabulated, two showed a .2 difference, ten showed a .1 difference, and two showed no difference whatsoever. The average quantitative ranking for the online format was 1.55; for the traditional format it was 1.46, with the highest possible rating 1.0 at "excellent", and 5.0 representing "poor". Three statistical tests (Independent *t*-test, Mann-Whitney, and Levine test for unequal variance) found no difference of statistical significance at the .05 level in these results. These statistical tests were administered because the data was based on group averages, rather than the individual scores of each student. The individual scores were not available, as results are reported to faculty only in a composite format. Students rated faculty essentially the same no matter whether the evaluations were online or in traditional form.

Extent of Open-Ended Responses: Although quantitative rankings can be extremely helpful in pointing out areas of relative strength and weakness, some instructors give more attention to open-ended comments. After completing the 14 Likert-scale questions, students in

the study were asked the open-ended question, “Do you have any additional comments?” In the area of open-ended responses, several differences were found between those who submitted online evaluations and those who submitted traditional evaluations.

Comments According to Format: Online responses to the open-ended question were consistently greater than were those from the traditional format. Of 670 individual comments, 452 were made online compared to 218 traditional—a ratio of 2:1. More than twice as many comments came from online evaluations as from traditional evaluations.

Comments within Returned Evaluations: Among those who returned evaluations, the percentage of online respondents making open-ended comments was also considerably greater. As Table 3 shows, when compared to the overall number of students who submitted evaluations in either form, approximately 27% more online respondents also wrote open-ended comments.

Table 3. Summary Data on Comments within Returned Evaluations

Students Submitting Evaluations (n=413)		Students Making Open-Ended Comments (n=250)		% Submitters Making Open-Ended Comments		
Online	Traditional	Online	Traditional	Online	Traditional	Difference
198	215	145	105	74%	47%	27% more from online

Comments and Words per Student: Despite the slightly higher return rate of traditional evaluations, of the 250 respondents who also took the time to make open-ended comments, online respondents made an average of half again (53% more) as many comments as their counterparts did. Furthermore, of the 7976 words written in the comments, the average number of online words was 54% greater than the average length of individual comments submitted in traditional form. Table 4 illustrates the differences.

Table 4. Summary Data on Comments and Words per Student

Students Making Open-Ended Comments (n=250)		Number of Comments (n=658) (Number per Student)		Number of Words Written (n=7976) (Number per Student)	
Online	Traditional	Online	Traditional	Online	Traditional
145	105	440 (3.03)	218 (2.08)	5395 (37)	2582 (24)

Nature of Comments (Positive or Negative): Comments made in the open-ended response portions of the evaluations were examined to determine whether there was a difference in the ratio of positive to negative comments in the two evaluation formats. As Table 5 shows, the

overall ratio of positive to negative comments revealed a 6% difference, with 55% of online comments being positive, compared to 61% of those submitted on traditional evaluations.

Table 5. Positive Comments vs. Negative Comments (N=644)

Online (n= 433)		Traditional (n=211)		Percent Online		Percent Traditional		Difference in Positive Comments
Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative	6% more from traditional
244	189	128	83	55%	44%	61%	39%	

Usefulness of Comments (Formative or Summative): After determining whether each open-ended comment was positive or negative, researchers categorized each comment as formative or summative, based on whether or not it provided the instructor with information on what to change, add or keep in the course. As shown in Table 6, online respondents made about twice as many statements, with online respondents making formative comments at a 6% higher rate than respondents submitting traditional evaluations (79% to 73%).

Table 6. Formative Comments vs. Summative Comments

Online (n=432)		Traditional (n=207)		% Online		% Traditional		Difference in Formative Comments
Formative	Summative	Formative	Summative	Formative	Summative	Formative	Summative	6% more from online
341	91	152	55	79%	21%	73%	27%	

Analysis

Overall, the results of this comparative study on course evaluation formats were these:

- It showed higher online return rates than other research had found, although there was still a somewhat higher return rate for traditional evaluations (7%);
- It supported other research on the similarity of quantitative rankings;
- It supported other research showing greater number and length in online responses;
- It showed differences in the area of open-ended comments and the extent to which these comments were positive or negative and formative or summative.

Return Rates. In response to concerns about return rates, this study differed somewhat from reports in the literature showing that online return rates were sometimes as low as 40% (Johnson, 2002). The results of this study showed a 76% return rate for online evaluations and an 83%

return rate for traditional evaluations. It is possible that the higher rates of return in this study were due to the use of a course management system (in this case, Blackboard); such systems have been found to increase response rates. Furthermore, it is possible that non-classroom settings provided the respondent with circumstances that are more amenable for some students, such as more time and more privacy from instructor and classroom peers. Although any number of factors can affect return rates, this study showed that comparable rates are indeed possible.

Quantitative Rankings. This is one of the most frequently voiced instructor concerns when choosing between traditional and online evaluation formats. This study verified what the literature points out consistently: there was no significant difference in quantitative rankings between online evaluation formats and traditional paper and pencil formats. This study found a slim .09 difference on a 5.0 scale.

Number and Length of Open-Ended Responses. As also reported in the literature, the number and length of open-ended responses was considerably greater in online evaluations than in traditional paper and pencil evaluations. The ratio of online comments to traditional comments was 2:1. The percentage of online respondents who made open-ended comments was 27% greater than the percentage of traditional evaluation respondents. The online respondents wrote 53% more comments per student than the traditional respondents (2.9 online to 1.9 traditional). Finally, online respondents wrote 54% longer comments than traditional respondents (37 words per student online to 24 words per student in traditional format).

Thus far the study seemed to show that online evaluations resulted in comparable return rates, comparable rankings, more open-ended comments, and lengthier open-ended comments. But what about the substance of those comments? In the seeming absence of previous research on this question, this study offers these findings.

Nature of Open-Ended Comments (Positive vs. Negative). Whereas quantitative rankings showed virtually no difference between the two formats, in the open-ended comments respondents submitting traditional evaluations showed a slightly higher percentage (6%). The researchers have not pursued this line of inquiry, but one reason may be the anonymity afforded by electronically printed comments as compared to handwritten comments.

Usefulness of Open-Ended Comments in Improving Instruction (Formative vs. Summative). Respondents completing online evaluations not only made more comments and longer comments, they made a larger percentage of formative comments (6%) than their counterparts did in traditional formats. Although a 6% difference may not seem large, when combined with substantially more online respondents making comments, and with their comments being substantially lengthier, the amount of information offered to instructors can be useful indeed.

Conclusions

Several conclusions can be drawn from this research, although it should be replicated in other settings with other populations. First, this study confirms existing research on similarities between evaluation forms. Second, it also clarifies what some faculty believe: There are indeed differences in student evaluations based on format. The differences, however, may allay rather than exacerbate faculty concerns. Based on our results, we offer the following for consideration:

1. Instructors using either format are likely to find that quantitative rankings will be similar;
2. Instructors using either format are likely to find that the proportion of positive to negative comments will be similar;
3. Instructors using online formats, however, are likely to find that open-ended comments will not only be quantitatively greater in number and length, but they will contain more qualitative detail than is likely to be found in traditional evaluations.

Faculty concerns are not to be dismissed lightly. Most educators would agree, however, that course evaluations are important not only for tenure and promotion but also to give faculty feedback to improve teaching, give students a chance for valuable input into their own learning, and help both strive for a learning community. All of these purposes seem well served by detailed, thoughtful student input into the class experience. If the findings of this study hold true, the combination of administrative convenience, instructor desire for timely feedback, and course improvement through formative evaluation may reduce the use of traditional paper evaluations and make online evaluations the format of choice.

References

- Anderson, H.M., Cain, J., & Bird, E. (2005). Online Course Evaluations: Review of Literature and a pilot study. *American Journal of Pharmaceutical Education*, 69(1), 34-43.
- Avery, R. J., Bryant, W. K., Mathios, A., Kang, H., Bell, D. (2006). Electronic Course Evaluations: Does an Online Delivery System Influence Student Evaluations? *Journal of Electronic Education*. 37:1, 21-38.
- Carini, R., Hayek, J., Kuh, G., Kennedy, J., Ouimet, J. (2003). College Student Responses to Web and Paper Surveys: Does Mode Matter? *Research in Higher Education*. 44:1; 1-19.
- Dommeyer, C, Baum, P., Hanna, R., Chapman, K. (2004). Gathering Faculty Teaching Evaluations by in-class and online surveys: Their effects on Response Rates and Evaluations. *Assessment and Evaluation in Higher Education*. 29:5, 611.
- Dommeyer, C, Baum, P., Chapman, K., Hanna, R. (2002). Attitudes of Business Faculty toward two Methods of Collecting Teacher Evaluations: Paper vs. Online. *Assessment and Evaluation in Higher Education*. 27:5, 455.
- Hmieleski, K., Champagne, M.V. (2000). Plugging in to Course Evaluations. *The Technology Source*. September/October. Available at:
<http://ts.mivu.org/default.asp?show=article&id=1034>.
- Johnson, T. (2002). Online Student Ratings: will Students respond? Paper presented at: American Educational Research Association Annual Meeting, 2002.
- Kasiar, J.B., Schroeder, S. L., Holstaad, S. G., (2001). Comparison of Traditional and Web-based Course Evaluation Processes in a required, team-taught pharmacotherapy course. *America Journal Pharmaceutical Education*. 63, 268-07.
- Kuhtman, M. (2004) Review of Online Student Ratings of Instruction. *College and University Journal*. 80:1, 64-67.
- Layne, B., DeCristoforo, J., McGinty, D. (1999). Electronic Versus Traditional Student Ratings of Instructions, *Research in Higher Education*. 40 (2). 221- 232.
- Oliver, R. & Sautter, E. (2005) Using Course Management Systems to Enhance the Value of Student Evaluations for Teaching. *Journal of Education for Business*. 80:4, 231-5.
- Ravelli, B. (2000). Anonymous Online Teaching Assessments: Preliminary Findings, Paper presented at: Annual National Conference of the American Association for Higher Education, June 14-18, 2000; Charlotte, North Carolina.
- Sorenson, D.L. & Johnson, D. (2003). *Online Student Ratings of Instructions*. San Francisco: Jossey Bass.

Wachtel, H.K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education*. 23(2), 191-211.

Appendix A Detailed Data Tables

Table 2. Data on Return Rates

Course Sections per Evaluation Format	Students		Evaluations Returned		Return Rate	
	Online	Traditional	Online	Traditional	Online	Traditional
1	28	27	19	23	.68	.85
2	27	27	24	17	.89	.63
3	25	15	15	10	.6	.67
4	15	28	12	22	.6	.79
5	28	29	21	26	.75	.9
6	22	19	17	14	.77	.74
7	28	30	19	25	.68	.83
8	18	16	10	16	.56	1.0
9	17	16	16	13	.94	.81
10	28	28	22	28	.79	1.0
11	25	23	23	21	.92	.91

Table 3. Data on Comments within Returned Evaluations

Paired Sections	Students Submitting Evaluations (n=413)		Students Making Open-Ended Comments (n=250)		Submitters Making Open-Ended Comments	
	Online	Traditional	Online	Traditional	Online	Traditional
1	19	23	12	6	.63	.26
2	24	17	16	9	.67	.53
3	15	10	12	6	.8	.60
4	12	22	10	9	.83	.41
5	21	26	13	23	.62	.88
6	17	14	10	5	.59	.36
7	19	25	15	22	.79	.88
8	10	16	8	2	.80	.13
9	16	13	12	3	.75	.23
10	22	28	15	7	.68	.25
11	23	21	22	13	.96	.62
Total	198	215	145	105	.74	.47

Table 4. Data on Comments and Words per Student

Students Making Open-Ended Comments		Word Count		Words per Student		
Online	Traditional	Online	Traditional	Online	Traditional	Difference
12	6	572	192	48	32	16
16	9	186	150	12	17	-5
12	6	341	148	28	25	4
10	9	405	106	41	12	29
13	23	413	669	32	29	3
10	5	225	101	23	20	2
15	22	596	556	40	25	14
8	2	281	29	35	15	21
12	3	272	44	23	15	8
15	7	1134	218	76	31	44
22	13	969	378	44	29	15

Table 5. Positive Comments vs. Negative Comments (N=644)

Paired Sections (n=22)	Online (n=433)		Traditional (n=211)	
	Positive	Negative	Positive	Negative
1	27	12	18	3
2	10	5	7	3
3	27	1	8	1
4	27	0	16	0
5	23	2	38	9
6	12	1	9	1
7	23	23	22	21
8	23	2	2	0
9	26	7	3	0
10	33	39	2	16
11	13	97	3	29
Total	244	189	128	83

Table 6. Formative Comments vs. Summative Comments (N=639)

Paired Sections	Online (n=432)		Traditional (207)	
	Formative	Summative	Formative	Summative
1	33	6	13	8
2	12	3	7	3
3	18	10	5	4
4	15	12	6	10
5	18	7	34	9
6	7	6	5	5
7	33	11	33	10
8	19	7	2	0
9	21	12	3	0
10	67	5	15	3
11	98	12	29	3
Totals	341	91	152	55
Percentages	79%	21%	73%	27%