

A Comparison of Web-Based and Paper-Based Course Evaluations at UNC Charlotte

*A Report Prepared by
The Center for Educational Measurement and Evaluation*

Richard Hartshorne, Ph.D.
Richard Lambert, Ph.D., Ed.S.
Bob Algozzine, Ph.D.

*And
The Center for Teaching and Learning*

Valorie McAlpin, Ed.D.
Mike Algozzine
Lee Norris
J. Garvey Pyke, Ed.D.

Table of Contents

EXECUTIVE SUMMARY 2

- Published Research Studies With UNC Charlotte Spring 2011 Comparison..... 2**
- Frequently Asked Questions 2**

BACKGROUND OF THE STUDY 3

- Purpose of Course Evaluations..... 3**
- Description of the Current Course Evaluation Process 3**
- Comparing Web-Based and Paper-Based Student Course Evaluations at UNC Charlotte 3**

METHODOLOGY 3

- Research Design 3**
- Participants 4**
- Instrumentation 4**
- Data Collection..... 4**

RESULTS..... 5

- Data Analysis 5**
- Comparison of Web-Based and Paper-Based Evaluations 5**
- An Evaluation of Efficiency of Resource Use for Paper and Web-Based Formats 6**

CONCLUSIONS..... 7

- Summary of Study Findings 7**
- Problems With Paper-Based Course Evaluations 8**
- Benefits of Web-Based Course Evaluations 8**
- Limitations of the Study 9**
- Final Recommendation..... 9**

BIBLIOGRAPHY.....10

APPENDIX A: MOTION TO FACULTY COUNCIL FROM FITSAC.....11

Executive Summary

Based on extant research, UNC Charlotte should consider the adoption of a university-wide, web-based course evaluation system for both formative and summative evaluations:

Ratings Similarity: Web and paper results show students very favorable towards instruction at UNC Charlotte. The results show no practical difference in ratings between survey methods and that open-ended responses are generally more complete in web surveys.

Availability of Results: Detailed web-based evaluation results are immediately available; this is a sharp contrast with the current system, where faculty wait for weeks or even months for results.

Data Reporting: A web-based course evaluation system provides opportunities for comparison of data across colleges, departments, and disciplines, thereby contributing to strategic planning and institutional research.

Response Rates: Prior studies show response rates in pilot tests of web-based course evaluations are typically lower than paper-based equivalents, but these rates rise when the web-based evaluations are centralized and fully operational with standard procedures in place. UNC Charlotte's pilot response rates significantly exceeded those in prior studies.

	Response Rate		
	Paper-Based	Web-Based, Pilot	Web-Based, Post-Pilot
Brigham Young	71%	50%	70%
Cornell	78%	50%	72%
Colorado State	75%	43%	n/a
McGill	55%	45%	51%
UNC Charlotte	92%	59%	n/a

Efficiency of Resource Use: Cost analysis indicates that a web-based course evaluation system would save at least \$218,000 annually, or \$1.09M over five years, a 73% reduction in total operating costs.

Frequently Asked Questions

Q. Are the results for web-based evaluations significantly different than paper evaluations?

A. No, there is no systematic tendency for web-based ratings to be higher or lower than paper ratings, even if the response rates are much lower. The ratings might be different, but higher scores are just as common as lower scores. When all the scores are combined, the positive and negative differences offset each other, resulting in web-based ratings that are not significantly or practically different than paper ratings.

Q. Is the response rate for online evaluations significantly lower than paper evaluations?

A. Sometimes, but typically only during the pilot phase or immediately after web-based evaluations are implemented. Once the system is fully operational and

standard procedures have been established, online response rates are equivalent to paper response rates.

Q. Is it possible to increase the response rate for online evaluations?

A. Yes, specific techniques can be used to boost response rates, including such things as instructor encouragement and standardized communication. Additional strategies can often be identified by analyzing specific factors that affect student participation at a particular institution.

Q. Are the results for online evaluations biased?

A. No, the results are not biased. The data collection method does not affect the results, and students with a strong opinion of the instructor are not more likely to participate. When variation occurs, it is due to factors other than the method of evaluation.

Background of the Study

Purpose of Course Evaluations

Student course evaluations at UNC Charlotte serve as a primary tool for formative and summative evaluations of faculty teaching. Resulting survey data are used by department chairs and deans in decision making related to annual reviews of faculty performance as well as related salary, reappointment, tenure, and promotion decisions. The existing decentralized student evaluation system currently in use at UNC Charlotte is paper-based and includes both required and optional survey items adapted from Purdue University's "cafeteria model." The student course evaluation system has traditionally been directed through the faculty governance structure, through the leadership of the Faculty Council.

Description of the Current Course Evaluation Process

Course evaluations are conducted via paper-based surveys distributed and completed during class time. Open-ended comments are included as appropriate or desired. The completed forms are then collected by administrative staff in each of the academic departments, who in turn hand deliver the forms to the OpScan office for scoring. Results are available via e-mail and printed hard copy, with hand-written, open-ended comments manually entered into the final report for each faculty member by administrative staff throughout the colleges. Student evaluations of fully online courses are typically conducted via a web-based enterprise system known as StudentVoice. Under both methods, evaluation results of faculty teaching are not typically shared with students. The current course evaluation process has served the needs of the university well until rapid growth during the last five years precipitated concerns over escalating costs and inefficiencies related to the economies of scale.

Comparing Web-Based and Paper-Based Student Course Evaluations at UNC Charlotte

To address the challenges associated with the ongoing implementation of paper-based course evaluations, UNC Charlotte began exploring the use of online student course evaluations as a replacement during the Fall 2010 semester. This project was designed to examine issues regarding the implementation of web-based student course evaluations, in comparison with the implementation of paper-based course evaluations. Areas of focus for this project included:

1. Determining if ratings are comparable between paper and electronic formats,
2. Analyzing qualitative feedback to determine if evaluation delivery medium impacts qualitative results,
3. Assessing response rates between paper and electronic evaluation formats, and
4. Evaluating efficiency of resource use between paper and electronic formats.

Methodology

Research Design

The research design for the study included eight course sections from each UNC Charlotte college—a stratified sample of two sections each from (A) small, lower-level courses, (B) large, lower-level courses, (C) upper-level undergraduate courses and (D) two graduate-level courses. Courses for inclusion in the study were recruited by the deans for each college at UNC Charlotte, who requested faculty volunteers to participate in the pilot study. Students in each section were randomly selected so that half of them completed the web-based course evaluation and the other half completed the paper-based course evaluation.

Participants

The research design sought participation from eight course sections from each college at UNC Charlotte (Architecture, Arts & Sciences, Business, Computing & Informatics, Education, Engineering, Health & Human Services), for a total of 56 sections. Due to logistical issues with two colleges, one college had only seven courses participate, and another college had only one course participate. Ultimately, the study consisted of 48 course sections. Of the eight course sections recruited from each college, two courses were small ($n < 30$) introductory undergraduate sections; two courses were large ($n \geq 30$) introductory undergraduate sections; two were upper level undergraduate sections ($n > 10$); and two were graduate sections ($n > 10$). Class sizes below 10 were not considered as they are considered exceptional and potentially different from other classes. The final sample included 1549 participants, with 774 assigned to complete the web-based student course evaluations and 775 assigned to complete the paper-based course evaluations.

Instrumentation

Prior to implementation of the paper-based and web-based student course evaluation study, participating colleges and departments submitted current copies of their existing course evaluation instruments to the research team. These were then converted to electronic formats via a third-party vendor, StudentVoice, for the web-based course evaluation group. While there were a few university required core evaluation items for each instrument, there was no common university-adopted instrument. In this study, no modifications were made to the questions or instruments submitted to the research team. However, after the study was completed, it was discovered that a number of departments did not submit open-ended survey items to the research team. Consequently, these open-ended items were not included in the study.

Data Collection

Each student in the participating course sections was randomly assigned to complete either the paper-based ($n = 775$) or the web-based course evaluation instrument ($n=774$). Those in the paper-based group did so in the usual framework of the course (e.g., during a class session near the end of the semester), while their peers completing the web-based evaluation did so via a website during a two-week window to complete it. To encourage participation, the web-based group received up to six e-mail reminders, each containing a link to the evaluation instrument; they did not receive reminder e-mails once they completed the survey.

In attempting to reconcile the data for analysis, it was found that evaluation forms used in each college varied a great deal and measured different aspects of teaching and learning. Thus, the researchers looked for common themes in the survey instruments and derived the following 12 domains for further analysis:

1. Overall satisfaction,
2. Learning effectiveness,
3. Instructional effectiveness,
4. Positive learning environment,
5. Varied instructional methods,
6. Use of instructional time,
7. Grading fairness,
8. Grading usefulness,
9. Instructor's preparedness,
10. Material relevance,
11. Course purpose, and
12. Instructor's availability.

Results

Data Analysis

Data analysis focused on comparisons between the paper-based and the web-based evaluations. Areas of analysis included the comparison of student ratings between paper and electronic formats, an examination of response rates between paper and electronic evaluation formats, and an evaluation of efficiency of resource use between paper-based and web-based formats. A full examination of open-ended, qualitative feedback was not completed at this time but is planned for the future.

Comparison of Web-Based and Paper-Based Evaluations

A total of 1,549 students were randomly assigned within the participating courses to complete their course evaluations in class using the paper-based process or to complete their course evaluations through the web-based system. ($n_{\text{Paper-based}} = 775$, $n_{\text{Web-based}} = 774$). A total of 1,171 students ($n_{\text{Paper-based}} = 714$, $n_{\text{Web-based}} = 457$) provided sufficient information to be included in the analysis. The response rate was very high for the paper-based condition (92.13%) and lower for the online condition (59.04%). A small number of students were dropped from the analyses ($n = 25$) due to incomplete data. At least five students responded in 39 different courses. One course was dropped from the analysis as only two students responded.

The comparisons between the two treatment conditions were completed by using multilevel modeling techniques. Students were nested within their courses. In the cases where there were multiple sections for a given course, the sections were combined. An average of 30 students responded per course (minimum=6, maximum=109). The level-one models, the within-course variance models, included the scale scores constructed from the course evaluation items as the dependent variables. A separate model was conducted for each outcome measure. Treatment group membership was entered as an uncentered predictor variable in the level-one models. The level-two models, the between-course models, were unconditional models with no predictor variables. Completely unconditional models were calculated as the first step in the analysis and 79.1% of the variance in course evaluation ratings was found to be within courses, while 20.1% of the variance in the ratings was between courses.

The scale scores were formed by taking the mean score across the respective items. As each item was rated on a five-point Likert-type scale, the scale score means were also measured on a 1-5 point scale with 5 being the most positive rating. As can be seen in Table 1, all of the group average ratings were above 4, reflecting positive ratings. There was a small, statistically significant difference ($p < 0.05$) between the treatment conditions for the rating of overall satisfaction. The paper-based group ratings were slightly higher than the web-based ratings; however, when expressed as a standardized mean difference effect size, this difference was small ($d = .16$). Students in both conditions were, on average, positive about the course experience. All scale score means, across both groups, were not lower than 4 on the 5-point scale. As shown in Table 1, a similar pattern of small, statistically significant differences was found for 9 of the 12 scale scores. For the remaining three scale scores, there was not a statistically significant difference between the groups. For none of the scale scores were the between group differences exceeding an effect size of approximately .21; coupling this finding with the possibility that the statistically significant differences were due in part to the large sample sizes in our analyses, we judged the practical value of all of the group differences to be small.

Table 1.

Difference Between Paper-Based and Web-Based Ratings by Scale Score.

Scale	Paper-Based Mean	Web-Based Mean	Pooled Within Group <i>SD</i>	<i>t</i>	<i>p</i>	95% CI Lower Limit	Effect Size	95% CI Upper Limit
Overall Satisfaction	4.424	4.330	0.585	-2.444	0.015	0.043	0.161	0.278
Learning Effectiveness	4.428	4.289	0.807	-2.725	0.007	0.055	0.172	0.290
Instructional Effectiveness	4.481	4.359	0.757	-2.346	0.019	0.043	0.161	0.279
Positive Learning Environment	4.517	4.421	0.748	-2.111	0.035	0.011	0.128	0.246
Varied Instructional Methods	4.469	4.348	0.777	-2.809	0.005	0.038	0.156	0.273
Use of Instructional Time	4.356	4.184	0.830	-2.845	0.005	0.089	0.207	0.325
Grading Fairness	4.310	4.250	0.787	-1.115	0.265	-0.041	0.076	0.194
Grading Usefulness	4.305	4.161	0.859	-2.375	0.018	0.050	0.168	0.285
Instructor's Preparedness	4.528	4.433	0.651	-1.577	0.115	0.028	0.146	0.263
Material Relevance	4.455	4.321	0.769	-2.753	0.006	0.057	0.174	0.292
Course Purpose	4.355	4.238	0.808	-2.210	0.027	0.027	0.145	0.262
Instructor's Availability	4.384	4.286	0.799	-1.647	0.100	0.005	0.123	0.240

An Evaluation of Efficiency of Resource Use for Paper-Based and Web-Based Formats

Web-based course evaluations generate substantial savings to the institution for materials and staff time (see Figure 1 below). Conservative estimates indicate that 80 hours of departmental staff time from each of 80 staff members is required to complete paper-based course evaluations with an annual cost of \$224,000 for personnel. Additional costs include \$15,000 for customized paper forms; \$5,000 in licensing costs for the existing web-based evaluation system currently used for distance education courses (i.e., this cost would be removed if the entire campus went to web-based student course evaluations); and \$6,246.40 in OPSCAN personnel costs. Therefore, the total annual cost is \$250,246.

The cost of licensing web-based course evaluation software for the entire university is \$24,500 annually. Coupled with the survey administration and management costs of \$56,500, UNC Charlotte would realize a cost savings of \$169,246, or a 68% savings in the operating costs of the student course evaluation process. A five-year savings of \$850M would be realized.

Figure 1. *Cost Savings of Web-Based vs. Paper-Based Student Course Evaluations*

Annual Paper-Based Evaluations Costs			
Description	Qty	Cost Per	Total
Cost of Paper Forms (including overprint)	100,000	\$ 0.15	\$ 15,000.00
Software Licensing Distance Education Web-based Course Evaluation Software	1	\$ 5,000.00	\$ 5,000.00
Departmental Staff Processing Time (80 Staff Members @ 80 hours each for manual processing of comments for all evaluations)	6,400	\$ 35.00	\$ 224,000.00
37% Reduction in OPSCAN Demand (See OPSCAN Utilization in Appendix)	488	\$ 12.80	\$ 6,246.40
ANNUAL COST OF PAPER -BASED EVALUATIONS:			\$ 250,246.40

Annual Web-Based Evaluation Cost Savings	
Description	Savings (Costs)
Paper Forms Savings	\$ 15,000.00
Reduction in OPSCAN Staffing	\$ 6,246.40
Distance Education Web-Based Evaluations License Savings	\$ 5,000.00
Staffing Productvity Savings	\$ 224,000.00
Annual Cost of Web-Based Software	\$ (24,500.00)
Administration and Management of Evaluations Across Institution	\$ (56,500.00)
Net Cost Savings:	\$ 169,246.40

Savings Summary	
Description	Savings (Costs)
Annual Paper-Based Evaluations Costs	\$ 250,246.40
Web-Based Cost Savings	\$ 169,246.40
Overall Reduction in Costs in %	68%
Five-Year Savings	\$ 846,232.00

Conclusions

Summary of Study Findings

Overall, it is clear from the results that students are rather favorable towards instruction at UNC Charlotte. As noted above, there were very small, statistically significant differences that slightly favored the paper-based student course evaluations. It is important to reiterate, however, that these differences were minimal. Furthermore, given the large sample size and the consistently low effect sizes, there is no practical significance in the difference in the ratings. In essence, because the domains selected typically

resulted in the favorable ratings noted above, these small differences across methods should not surprise administrators or faculty.

Problems With Paper-Based Course Evaluations

The paper-based course evaluation method has several limitations:

- Escalating institutional costs for paper, printing, distribution, collection, and storage.
- Transcribing hand written comments allows for subjective interpretation of data and additional resources of staff time.
- Administering evaluations in the classroom setting limits the amount of time students are able to dedicate to the evaluations.
- Administering evaluations in class requires devoting a portion of class time to completing evaluations.
- The classroom setting poses limitations on the effectiveness of the evaluations; students complain of being unable to contribute thoughtful comments in a short timeframe.

Additionally, the existing UNC Charlotte decentralized student evaluation system lacks uniform administrative support, which makes university-wide data comparisons of faculty teaching difficult and unwieldy. Provisions for administrative oversight, support, and coordination have not been considered.

Benefits of Web-Based Course Evaluations

The web-based course evaluation method has several benefits to faculty, students, and the institution:

- Lower turnaround time to deliver immediate feedback to faculty, department chairs, and deans.
- Increased ability to perform statistical analyses with course evaluation data.
- More substantive feedback from students on open-ended questions.
- Increased efficiency, less manual manipulation required by administrative staff.
- Better data, since errors are less likely and open-ended responses are generally more complete.
- Universal access for all students, regardless of class attendance.
- Substantial savings to the institution for materials and staff time, including reduced printing, distribution, collection, and storage costs.

While a detailed quantitative and qualitative analysis of the open-ended responses is ongoing, a cursory review of these responses indicated that there was a significant increase in the quantity of open-ended responses on the web-based student course evaluations. This was even more significant, as a number of the participating departments omitted the open-ended responses from their evaluation instruments. This cursory finding aligns with previous research that cites additional time as a key indicator of both quality and quantity of open-ended responses. Studies at other universities show that errors are less likely and open-ended responses are generally more complete. Consequently, web-based evaluations typically generate better data than paper.

According to *Online Student Course Evaluations: Strategies for Increasing Student Participation Rates* (2009), a custom research brief prepared by the Education Advisory Board, the greatest challenge in converting to a web-based course evaluation system is the decline in student response rates that institutions often experience during the first year of transition. With a centrally supported, controlled environment in which to administer course evaluations, student response rates generally return in year two to the pre-online rates.

Limitations of the Study

There are several potential confounding issues related to the response rate in the study. First, the response rate difference between treatments may be a function of the experimental design of the study, and it disappears when we only offer one option. With that said, we believe that the research design selected (i.e., randomly assigning participants within courses to each group rather than selecting entire courses to complete either the online student course evaluation or face-to-face student course evaluation) was more rigorous and provided us with more powerful results.

Secondly, a number of faculty participants cited confusion with the selection of the web-based participants (e.g., students were not sure if they received the e-mails, etc.). This had a compounded effect on the response rates in the study, as faculty noted the possibility of confused students accidentally completing the paper-based course evaluations, even though they were in the group designated to complete the online student course evaluations.

Students were likewise confused by receiving email from StudentVoice to notify or remind them to complete the web-based evaluation. Since students were not familiar with StudentVoice, many of them treated these email messages as spam and likely never completed the evaluation. This could have had a significant impact on response rate, since the emails did not come directly from UNC Charlotte.

One final limitation of the study is the lack of a common course evaluation instrument at UNC Charlotte. In attempting to reconcile the data for analysis, it was obvious that student course evaluations from each college varied a great deal and were designed to measure very different aspects of teaching and learning. Thus, the researchers had to derive common themes in the survey instruments and classified them into the 12 domains cited in the study.

Final Recommendation

UNC Charlotte should consider the adoption of a university-wide, web-based course evaluation system for both formative and summative evaluations. Access to course evaluation data in a timely manner will allow faculty to make course revisions as needed for greater student success and improved learning. Additionally, administration of a centralized course evaluation system is encouraged to allow for greater comparison of data across colleges, departments, and disciplines, economies of scale, and increased access to reporting features.

Bibliography

- 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.
- Dommeyer, C.J., Baum, P., Hanna, R.W., & Chapman, K.S. (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), 611-623.
- Johnson, T. D. (2003). Online student ratings: Will students respond? *New directions for teaching and learning: Online student ratings of instruction*, 96(49-59).
- Layne, B.H., DeCristoforo, J.R., & McGinty, D. (1999). Electronic versus traditional student ratings of instruction. *Research in Higher Education*, 40(2), 221-232.
- Winer, L.R. & Sehgal, R. (2006). *Online Course Evaluation Analysis Report*.
http://www.mcgill.ca/files/tls/online_course_evaluation_report.pdf
- The Office of University Planning and Analysis (UPA) at North Carolina State University
<http://www2.acs.ncsu.edu/UPA/assmt/resource.htm>
- OnSET: Online Student Evaluation of Teaching in Higher Education
http://onset.byu.edu/index.php?title=Main_Page

Appendix A: Motion to Faculty Council From FITSAC

Memorandum

August 17, 2011

TO: Faculty Executive Committee

FR: Heather McCullough, FITSAC chair

RE: Motion to implement web-based student course evaluations

Motion

FITSAC has reviewed the findings of an internal study approved by the Faculty Council and conducted by the Center for Educational Measurement and Evaluation (CEME) and recommends the adoption of a university-wide, web-based student course evaluation system.

Rationale

A research team conducted a comparative study of paper-based versus web-based evaluations that included 1,549 student respondents, representing each college at the university and a range of graduate and undergraduate courses. The study found that there were no practical differences between the methods of evaluation—web-based or paper-based—in the reported level of student satisfaction with their courses and instructors. Of the domains of question types used in the evaluations, small, statistically significant differences between the paper-based and web-based conditions were found for 9 of the 12 scale scores but the practical value of all of the group differences was found to be minimal.

In addition to the results that student responses were not affected by the evaluation medium, the study indicates significant resource efficiencies in personnel, paper, printing, and equipment costs with the added benefits of immediate access to data. Conservative estimates indicate a savings of \$218,000 per year, a 73% reduction over existing costs.