

Online and In-person Evaluations: A Literature Review and Exploratory Comparison

Paulette Laubsch

Assistant Professor

School of Administrative Science

Fairleigh Dickinson University

Teaneck, NJ 07666

plaubsch@fd.edu

Abstract

There has been rapid expansion of online degree programs over the past decade; with increases of asynchronous computer based instruction tripling between 1995 and 1997 and similar increases reported through the 2001 reporting period by the National Center for Education Statistics (nces.ed.gov). Many institutions of higher education have been philosophically and technically struggling with if and how to transition some, if not all, in-person classes and processes into an appropriate online modality. Skeptics of online coursework have challenged the ability of programs to maintain the same level of quality as the in-person classes. Student evaluations of the class and course instructor are one tool to assess the quality of courses and faculty. This article compares in-person and online evaluations for adjunct faculty in one graduate program.

Introduction

Student evaluations of faculty tend to be reliable measures of quality, although instructors who are enthusiastic or have positive personality traits generally fare better on student evaluations (Obenchain, Abernathy, & Wiest, 2001). Waters, Kemp, and Pucci (1988) concluded that instructors are rated higher by students if they have what the authors refer to as favorable personality traits. Radmacher and Martin (2001) found a significant relationship between faculty extroversion and the students' perceptions of teaching effectiveness. The move to an online modality removes many of these personality traits from the process; particularly, non-verbal communication and use of humor and other types of communication. In an early study of evaluations, Bean (1978) found that student evaluations could be influenced by others surrounding them in the evaluation setting. This social influence is eliminated in the online modality since students generally complete the process on their own.

With regard to matters such as tenure and promotion, accreditation, and faculty development, faculty evaluations are stressed, albeit at different universities and different disciplines to different degrees, as a means of ensuring accountability, and this is not a new phenomenon or strictly limited to one area or nation (Ballantyne, 2003). Faculty evaluations completed by students are used as partial criteria for promotional consideration and tenure, but the weight of such evaluations varies across institutions (Chang, 2003; Obenchain, Abernathy, & Wiest, 2001; Radmacher & Martin, 2001). What further complicates the issue is the fact that more institutions of higher education are utilizing higher numbers of adjunct faculty (Ehrenberg and Zhang, 2004) who do not qualify for promotion or tenure and may not be subject to the same evaluation process for any number of reasons, including, but not limited to, lack of administrative staffing to

monitor such a process, particularly in-class observations.. This results in a large number of individuals not being evaluated and possibly not being held accountable. This issue of evaluation may also relate to motivation on the part of the instructor with potential differences between adjunct faculty with different long term career visions, and junior and senior faculty, however, this is outside the scope of this article.

Focus of this Article

The purpose is to determine whether there is a difference in quality of instruction between the two modalities as reported by students. This paper is not meant to assess the validity of student evaluations of instructional staff nor is it designed to assess the impact of evaluations on instructional quality. The author compared in-class evaluations with those for online coursework for adjunct instructors. This article reports on a preliminary study of course evaluations at one graduate level university program. Students in the program generally come from the public and not-for-profit sectors where evaluations are an expected protocol, and students and faculty are aware of this at the onset of respective courses.

Online Evaluations

A study by Carini, Hayek, et al. (2003) revealed only small differences between online and in-person paper evaluations. In their study, nearly 90% of the students were under thirty years of age. There may be differences when teaching older, particularly more seasoned future scholar-practitioners such as in an applied graduate program as is discussed herein. Another factor could be higher levels of responses from students at “highly wired” campuses than other institutions. There has been some research conducted on response rates, with initial hypotheses stating a predicted lower response rate for online instruction; presumably because this is at the convenience of the student, as contrasted with in-person, on-site distribution (Ballantyne, 2003; Dommeyer, Baum, & Hanna, 2002). Faculty support and promotion of online processes, including encouraging completion of the process and explaining the importance of such evaluations, has indicated higher numbers of responses (Ballantyne, 2003). Research on undergraduate courses and evaluations in an online mode has also been undertaken (Chang, 2003; Dommeyer, Baum, & Hanna, 2002). In this research, Chang found that in-person evaluations were higher in ranking than online evaluations.

Although there may be many issues and concerns with use of online evaluation formats, sufficient information is available to show they are a valid methodology. Although maintaining an online rating system requires administrative support (Goodman & Campbell, 1999), there are a number of advantages of online collection of data compared to paper formats, including, but not limited to, less class time use, faster data processing, and the ability to get more in depth information (Dommeyer et al., 2004). Costs associated with the traditional paper survey are many: printing, distributing, collecting, scanning or tabulating, typing open-ended student responses, and delivering hard copies of the summaries to faculty and administration (Dommeyer et al., 2004). Students who are comfortable with the Internet and use of web sites should have little difficulty in completing any online evaluation form, but they may feel that it is inconvenient for them to do so. Absent the “pressure” of in-class evaluations, this also raises issues in terms of potential skewing of electronic responses, perhaps only for students who were dissatisfied for some reason.

Response Rates for Online Surveys

As online learning has continued to grow (Carini et al., 2003), more institutions are looking at a better model for faculty evaluations. Although there have been some innovations in online faculty assessment processes (Mandernach, 2005), standardized faculty evaluation formats are still being used, despite issues of emphases endemic to some disciplines. A primary concern is the response rate. An in-class evaluation should have close to 100% of the students attending that particular class session responding. There may be a few students who do not want to fill out the

form, but when done at the end of the course, most students complete the form, although they may not respond completely to open-ended questions. Also, with in-class evaluations that have a rather broad time frame for completion in class, a possibility is raised that an instructor could withhold implementing the process until a known and dissatisfied student is not present.

In 1999, high level of computer literacy and access to computer equipment for 24 hours per day were needed for one school to initiate a pilot program of an online evaluation system for students enrolled in in-person classes (Ballantyne, 2003). The online process resulted in 30% response rate compared to the 65% response rate for the paper surveys. This raises the issue of reliability of the responses (Ballantyne, 2003). Various researchers have provided parameters of the percentage of responses needed for a valid response. Dommeyer, et al. (2002) cited Centra, who stated at least two-thirds of the student responses were needed, and Gilmore, Kane, and Naccarato, who argued that a smaller number could be acceptable if the course were taught multiple times during the period of faculty review

When faculty actively promote completion of online surveys, there is an increase response (Dommeyer et al, 2004; Ballantyne, 2003). In the pilot process reported by Ballantyne (2003), responses rose to 54% and then 72% in 2000. The following two years, the response rates dropped to the 50% level. Ballantyne also reported on the use of online evaluations for online classes that provided wide ranging response rates. Rates for paper surveys were approximately 40% in one school, and when online surveys were implemented, the rates stayed the same initially. It was determined that there were significant response rates for online courses, with the rates varying between 30 and 95 percent.

Student Concerns about the Process of Course and Instructor Evaluations

Students have indicated a preference for the online format (Layne et al., 1999; Dommeyer et al., 2004), although they think the process takes too long. Part of the length issue is related to responses to the open-ended questions. Online surveys result in more open-ended responses since there is additional time available (Dommeyer, Baum, Chapman, & Hanna, 2002). Issues of anonymity are also a concern of the students. To overcome this issue, there needs to be a statement provided that guarantees the students anonymity and confidentiality of their response (Dommeyer et al., 2004; Ballantyne, 2003; Carini et al., 2003).

In many classes, the in-person evaluation forms are completed during the last class. This is prior to the final grade being computed. Online evaluations are accessible to the students for a particular period of time and allow for students to complete the form over that period of time (Dommeyer et al., 2004), but those that can be completed after the end of the class may be impacted by grades received, a reward or punishment factor that may influence the individual completing the form. This may result in a validity and possibly a reliability issue since students who receive a lower than anticipated grade may rate the instructor and course lower or the individual receiving a higher grade rating the instructor higher. Although the use of online evaluations for in-person classes has increased (Hoffman, 2003), such methods have provided mixed results (Carini et al., 2003). Later studies have indicated that the evaluations are reliable and produce higher quality and greater quantity of response to open-ended questions (Dommeyer, Baum, Hanna, & Chapman, 2004).

Online versus In Class Surveys

Analyses of in-class teaching evaluations compared to online surveys appear in various studies. Dommeyer, Baum, Hanna, and Chapman (October 2004) conducted a study in 2000 comparing faculty attitudes towards evaluations completed in class and those collected online. Online evaluations generally utilized the Internet, a web site address that can be hosted by the institution or even a third party where the survey instrument could be found. Although ideally information contained in these surveys should only be entered by the students enrolled in the class and

should only be entered once, use of the Internet for this process utilizes a student identification number without other authentication (Ballantyne, 2003) or even a user identification and password. Individuals who are less than satisfied with the instructor or the class may be able to respond more frequently, which results in a negative review. Information submitted on the surveys is provided to instructors in aggregate form to further allow for anonymity.

Ballantyne (2003) questions whether the flexibility of course format should also result in a flexible evaluation process. Related to this is the degree of flexibility allowed: should the student be allowed to complete the evaluation in-person or online at any particular time depending on their own schedule and desire. Administrative issues arise such as how to ensure non-duplication of assessments and how to ensure the individual completing the survey is truly the individual in the class.

Other studies have been conducted on online course evaluations, including work by Hmielecki, as reported by Hoffman (2004), surveyed 200 of the most wired institutions in 2000 and found that 90% of the institutions still utilized paper survey instruments for student evaluations. Hoffman reported on a follow-up survey completed between October 2002 and December 2002 with a sample of 500 institutions. The results of the later survey indicated approximate 10% of the institutions used an Internet system that was used throughout the school's programs as a principal means for collecting evaluations. Although the percentage remained the same (10%), the institutions were not only the most wired schools.

Studies of demographic characteristics and responses to paper versus online surveys have generally been limited to undergraduate students. Males who are more affluent and younger are more likely to complete online surveys than their female counterparts (Sax, Gilmartin, & Bryant, 2003). Sax et al. (2003) found many additional concerns with online modalities. Lengthy online survey instruments have fewer responses than shorter ones. Students are exposed to more emails and demands on their online resources and may be at the saturation point when surveys are to be completed. This study included traditional college students, who are more inclined to complete online processes.

There have been some studies on personal characteristics and their influence on student evaluations of faculty (Lumsden and Scott, 1995). In a classroom setting, faculty may influence how students perceive them by rewarding them or providing them with some other incentive, but in an online environment, many of the personal characteristics are negated (Dommeyer et al., 2004). Since students are able to complete the survey at their convenience and from their own computers, there is a distance between faculty and the process. Faculty also do not handle the forms and thus have no opportunity to manipulate ratings through comments when survey forms are distributed or to modify evaluations prior to submission (Dommeyer, Baum, Chapman, & Hanna, 2002).

Description of In-person, Blended, and Online Instruction Used in the Study

The Master of Administrative Science (MAS) degree program at Fairleigh Dickinson University is offered at selected off-campus sites throughout the State of New Jersey, as well as online. This program is applied in nature with admission primarily contingent on academic history and practitioner status, as a general rule, but with exceptions. Thus it primarily serves mid-career adults, generally in the public and not-for-profit sectors. As a group, these students may not have the same level of technical expertise with regard to use of computers as undergraduate students who have been socialized into technology and which may impact their acceptance of technology.

Courses are offered in a trimester format with classes held one night a week for twelve weeks, on five consecutive Saturdays, or one week long seminars. In addition to the in-person format, courses are offered in a blended model, as well as fully online courses. Sufficient courses have been developed in an online modality to allow students to complete their degree through the

distance learning curriculum. In addition, class size is maintained at a relatively low number. The table below indicates average class size during the past two years.

Term	Courses	Enrollment	Average Class Size
SP 04	6	87	14.5
S1 04	8	114	14.2
FA 04	17	214	12.6
SP 05	20	301	15.0
S1 05	23	311	13.5
FA 05	28	425	15.2

Average Class Size rounded to nearest tenth

As the program grows each year, there is an increased demand for faculty to facilitate courses. Once adjunct faculty expresses interest in such assignments and gains competence in this role, the program monitors the quality of instruction and also responds to market demand.

Since coursework is developed to meet emerging trends and student demand, instructional staff need to possess both employment experience and depth of knowledge in many areas; again, the program is an applied master's degree. While programs that have attempted to utilize full-time faculty who lacked such depth of experience and knowledge found their instructors to fare poorly in classroom situations (Lumsden and Scott, 1995), the MAS program seeks individuals with the requisite experience and knowledge as well as educational credential. Many of these individuals are employed as adjunct faculty, who are recruited and employed based on their expertise in a specific area of specialization. They lack training in educational theory or methodologies. Adjuncts are mentored by full-time faculty in instructional theory. Full-time faculty have developed guidelines for syllabi; methodologies for teaching adults; provided training sessions all in various university processes; and engaged adjuncts in developing appropriate coursework.

At the beginning of the distance learning initiative, faculty with specific expertise were asked to transition in-person course content into a web-based course, with the assistance of one on one technical support personnel. Specific guidelines were established relative to number of learning units, instructional/lecture notes, assessment tools, and discussion questions, in addition to the standardized syllabus guidelines that must be followed for every course. Once a course is developed, program administrators review the course to ensure it meets the established guidelines. Individuals who develop courses are generally given the option of facilitating the course when it is scheduled.

In an effort to facilitate online course development and usage, full-time faculty in this one program have undertaken the mentoring and development of adjunct faculty for the online coursework. Individuals are provided guidelines for course development and receive practical training in the operation of Blackboard, the web-based application used for delivery of coursework. These individuals are taught how to operate the Blackboard system in the initial training that provides the basics, such as how to sign on, send emails, maintain course documents, and develop discussion questions. During the second term, additional Blackboard functions are reviewed along with any specific issues that had arisen during the previous term. Significant effort was spent to ensure the same level of integrity for the inputs of the online and in-person classes. The program also has focused attention on the end of course assessments as a measure to assess outputs in both modalities; these are operationalized herein for the purpose of this article as course and instructor evaluations.

Student evaluations of the instructor, the facility, books, and other aspects of the courses have been present from the inception of the program. The MAS program utilizes its own evaluation form for adjunct faculty. The University has a standardized evaluation format for full-time faculty. This paper survey for full-time faculty has been converted to an online evaluation format for the MAS program only and is accessed through a Web-based application. The development of

online classes provided a unique opportunity to modify the evaluation tool to address issues relative to the quality of the materials as opposed to personality factors for adjunct faculty. As a non-traditional program operating off-campus, the School of Administrative has considerable latitude in administrative procedures. Since the faculty are mainly adjunct, development of the online evaluation process has been possible without the usual administrative issues (Dommeyer et al., 2002). In addition, questions were added to address specifically the technology issues, such as difficulty in accessing the system or connectivity problems. Evaluations are reviewed by the Dean and the program's academic coordinator, and appropriate action is taken in response to issues raised. This minimizes issues with students, and helps ensure a good-quality product for the students. Evaluation of instructors' performance from an administrative view is another topic that needs to be explored. However, this study only looks at student evaluations of instructors.

Methodology

This study involved a comparative analysis of categories of responses to selected items. The instruments used for in-person and online evaluations differ since they measure aspects relative to the modality of presentation; thus, the evaluations could not be completely compared.

In the MAS program, the in-person classes have close to a 100% return rate for evaluations since they are distributed at the last class and collected by one of the students, who seals the forms in an envelope that is returned to the administrator's office. Support staff summarize the individual responses in the numerical part of the evaluation form and type the open-ended responses. The summary forms with the comments are sent to the instructors, and a copy is maintained in the office files.

Online course evaluations are reserved for online classes only and have a format that is activated during the last week of the class by the program administrator. At the end of the term, the support staff and the instructor can access summary information and print such information for his or her records. This includes the open-ended responses.

For the fall 2005 term, 192 online evaluations were completed for the 315 students enrolled in courses taught by adjunct faculty, for a return rate of 61.0%. The in-person classes had 310 evaluations returned for 358 students enrolled, however only 294 were complete and used in this study. This equates to a return rate of 82.1%. Some students may be absent when the forms are distributed or do not complete the form for some reason.

A Likert scale was used in the evaluation process and ranged from 1 for strongly agree to 5 for strongly disagree. The responses were weighted, using the Likert scale. Since students could elect to not respond to any of the questions in either format, each question was scored based on the number of actual responses. The comparable questions related to the following:

- A. Materials were clearly presented
- B. The course was well structured
- C. Assignments reinforced materials in the course
- D. The course materials were perceived as relevant to the students

The weighted averages and standard deviations are shown in Table 1 below.

	Number of Responses	Question A	Question B	Question C	Question D
In-Person	294	1.20 ± .24	1.22 ± .23	1.20 ± .21	1.42 ± .32
Online	192	1.56 ± .41	1.64 ± .51	1.47 ± .28	1.64 ± .42

Table 1. Comparison of Selected Questions Using Means and Standard Deviation

Two instructors taught both online and in-person classes. Instructor X taught three classes: two in the online modality and one in-person. One online class had 5 students, and the other had 12. Aggregated data were used for the analysis. The average scores for the four questions are provided in Table 2 along with the standard deviations.

Instructor	Modality	Responses	Question A	Question B	Question C	Question D
X	In-Person	12	1.25 ± .13	1.33 ± .13	1.25 ± .13	1.25 ± .13
	Online	17	1.88 ± .21	1.76 ± .21	1.47 ± .21	1.47 ± .21
Y	In-Person	14	1.21 ± .29	1.50 ± .29	1.29 ± .29	2.14 ± .29
	Online	14	1.21 ± .05	1.29 ± .05	1.29 ± .05	1.21 ± .05

Table 2. Comparison of Instructor Teaching in the Two Modalities

The two instructors who taught in both modalities had different results. Although personality might have influenced the in-person scoring, Instructor Y was fairly consistent in both modalities. In attempting to determine reasons for the differences between the two instructor scores, it was found that Instructor X taught different subjects but Instructor Y taught the same subject.

In addition to the assessment questions that are graded using a Likert scale, both assessment instruments provide for open-ended questions. Each of the tools had three questions. The number of comments for each of survey was recorded as was the total word count. Since these narrative responses were entered into a word program, the tools option for word count was used to determine the total number of words entered by the respondents in both formats. The total averages and standard deviations are shown in Table 3.

	Number of Students	Question	Total Comments	Comments Per Student	Total Words	Words Per Comment
Online	192	#1	168	.88	3490	20.8
		#2	161	.84	2388	14.8
		#3	120	.62	2935	24.5
		TOTAL	449		8813	
		Mean		.78 ± .14		20.03 ± 4.90
In-Person	294	#1	172	.59	1610	9.4
		#2	97	.33	829	8.5
		#3	102	.35	1493	14.6
		TOTAL	371		3932	
		Mean		.42 ± .15		10.83 ± 3.29

Table 3. Summary of Comments and Words Per Comment in Online and In-Person Evaluations

Analysis and Findings

This study looked at one particular program and one group of instructors, the adjunct faculty, in two modalities, online and in-person delivery of course materials. The purpose was to determine if there were differences in student evaluations related to the instructional mode. There were 20 online courses and 29 in-person courses. The scores are ordinal numbers.

Since only the data for Questions A, B, and D appear to be normal distributions, a non-parametric test was used for data analysis. Non-parametric tests are used whenever assumptions on data normality or variances are not made. SPSS 12.0 for Windows was used to for the Mann-Whitney Test (Table 4) to compare medians. For each Question A through D, the null hypothesis, there is no difference between the median online responses and the median in-person responses, was

tested. The alternate hypothesis is the median response to the online course is greater than the median response for the in-person course for each question. The level of significance, alpha, is 0.05. The 95% Confidence Interval of the Difference for each question and each modality is also provided. This statistic indicates that 95% of the time, the scores should fall between the lower and upper range. The column titled Test of Significance is the results of the test of online responses equals in-person responses compared to online responses being greater than in-person responses.

Question		n	Median	Point Estimate for Change	Test of Significance	95% Confidence Interval of the Difference	
						Lower	Upper
A	Online	20	1.45	.33	.0003	.13	.50
	In-person	29	1.12				
B	Online	20	1.54	.29	.0003	.13	.53
	In-person	29	1.14				
C	Online	20	1.45	.25	.0001	.12	.40
	In-person	29	1.17				
D	Online	20	1.58	.15	.0321	.00	.42
	In-person	29	1.29				

Table 4. Mann-Whitney Test Statistics for Questions A through D

For Question A, the test of significance result of .0003 is less than 0.05. Therefore, the null hypothesis is rejected, and the conclusion is that there is a significant difference between the two group responses to the question. Therefore, the online responses were greater than the in-person responses at the 0.05 level of significance.

Similarly, for Question B, C, and D, the test of significance for each is less than the 0.05 level of significance, alpha. The null hypotheses are rejected, and the online responses are greater than the in-person responses at that 0.05 level.

If the hypothesis for testing comments and word counts is that there is a difference between the online and in-person groups, the null hypothesis is that there is a no difference between these groups. The average counts were used in a t-test to analyze the data, the results of which are shown in Table 5.

		Test Value = 0					
		t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
						Lower	Upper
Comments	Online	9.650	2	.01	.78	.43	1.13
	In-person	5.068	2	.04	.42	.06	.78
Word Count	Online	7.088	2	.02	20.03	7.87	32.19
	In-person	5.698	2	.03	10.83	2.65	19.01

Table 5. T-test Data for Comments and Word Count

Looking at the observed two-tailed significance level, one can expect to see the sample differences 1%, 4%, 2%, and 3% of the time if the samples were equal. Each of these values is less than 5%, which rejects the null hypotheses for each. Therefore, there are differences between the group.

The following summarize the findings of this study:

- There is a difference in response rates with lower percent of online responses compared to in-person responses, at a level of significance of 0.05.
- The in-person scores indicate that the students more strongly agree with the items included in the surveys than do the online students.
- Neither set of scores were what program administrators could consider poor ratings since each online issue rated better than agree.
- When comparing online and in-person responses for the same instructors, one exhibited similar scores and the other had a difference similar to the overall ratings.
- There are increased numbers of open-ended responses in online surveys.
- Online open-ended responses result in more responses and greater quantity of information.

These responses rates may have more to do with issues involving delivery of the subject and personality factors of the instructor than with the subjects themselves. As stated in the literature, many evaluations are influenced by personality factors, which do not necessarily appear in the online mode.

There may be other factors that result in higher in-person scores. The students may anticipate that the instructor would know what they reported for each of the categories and possibly fear an impact on their grades. Although care has been taken to reduce the potential of instructor manipulation of the process, there may be some influence over the ratings by the instructors.

Students may also be affected by the class situation where they are rushed to complete the form so that they do not infringe on class time, or in some cases, the time for the exam. In an effort to rush through the process, the students might just check a box without completely assessing the situation.

The open-ended responses affirm the findings of Dommeyer, Baum, Chapman, and Hanna (2002) as to higher quantity of responses from online instruments. Although there is a lower response rate, the number of responses and the number of words per response is much higher than the in-person surveys. This may be related to the time factor since the in-person responses are confined to the start of the class and thus restricted to the time frame established by the instructor. The online assessment does not have such a restriction and allows for more responses. No assessment was completed on the quality of the responses, but the increase in the number of words in the online mode might be an indicator of more in depth responses, and therefore of higher quality.

Although the difference in response rates between the two modalities is large (61.0% for online compared to 82.1% for in-person) and the averages are higher for the online than the in-person, the online process may yield better assessments than the in-person process since students have time to be reflective about the course and take as much time to complete the format as needed or desired. This type of reflective thought and response may be useful for program administrators responsible for assessing faculty. Programs should consider higher utilization of online assessment tools for assessing faculty and program offerings.

The student population for this program is generally the mid-career professional. It is possible that response rates for other groups would be higher based on their computer expertise.

Additional studies should be undertaken to determine if there is a correlation between age, computer expertise, or course of study and response rates as well as average scores for faculty and despite institutional protectiveness; across institutions or at minimum, comparison of literature as pertains to specific institutions.

Special Acknowledgment to Dr. Lisa Marano, Assistant Professor of Mathematics, West Chester University. Dr. Marano provided technical assistance for statistical methodology and analysis.

References:

- Ballantyne, C. (Winter 2003). New directions for teaching and learning. Wiley Periodicals, Inc., 96, pp. 103-112.
- Bean, L. H. (December 1978). Effects of the social circumstances surrounding the rating situation on students' evaluations of faculty. *Teaching of Psychology* (5)4, pp. 200-202.
- Carini, R. M., Hayek, J. C., Kuh, G. D., Kennedy, J. M., & Ouimet, J. A. (February 2003), College student responses to web and paper surveys. *Research in Higher Education*, (44)1, pp. 1-19.
- Chang, T. (April 21-25, 2003). The results of student ratings: The comparison between paper and online survey. (Article presented at American Education Research Association, Chicago, Illinois meeting held April 21-25, 2003).
- Dommeyer, C.J., Baum, P., Hanna, R. W., & Chapman, K. S. (October 2004). Gathering faculty teaching evaluations by in-class and online surveys. *Assessment & Evaluation in Higher Education*, 29(4), pp 611-623.
- Dommeyer, C. J., Baum, P., Chapman, K. S. & Hanna, R. W. (2002). Attitudes of business faculty towards two methods of collecting teaching evaluations: Paper vs. online. *Assessment & Evaluation in Higher Education* (27)5, pp. 455-462.
- Dommeyer, C.J., Baum, P., & Hanna, R. (September/October 2002). College students' attitudes towards methods of collecting teaching evaluations: In-class versus online. *Journal of Education for Business*, (78)1, pp. 11-15.
- Ehrenberg, R. G. and Zhang, L. (2004). The Changing Nature of Faculty Employment. (Prepared for the TIAA-CREF Institute Conference *Retirement, Retention and Recruitment: The Three R's of Higher Education in the 21st Century*, presentation in New York City, April 1-2, 2004).
- Goodman, A. and Campbell, M. (1999). Developing appropriate administrative support for online teaching with an online unit evaluation system. *Advances in Multi-media and Distance Education*, pp. 17-22.
- Hoffman, K. M. (Winter 2003). Online course evaluations and reporting in higher education. Wiley Periodicals, Inc., 96, pp. 25-29.
- Layne, B. H., DeCristoforo, J. R., & McGinty, D. (1999). Electronic versus traditional student ratings of instruction. *Research in Higher Education*, (40)2, pp. 221-232.
- Lumsden, K. and Scott, A. (April 1995). Evaluating faculty performance on executive programmes. *Education Economics*,(3)1. pp. 19-31.

Mandernach, B. J. (Fall 2005). A faculty evaluation model for online instructors: Mentoring and evaluation in the online classroom. *Online Journal of Distance Learning Administration*, (8)3.

National Center for Educational Statistics (NCES) (retrieved from http://nces.ed.gov/programs/quarterly/vol_2/2_1/q5-7.asp, February 15, 2006)

Obenchain, K. M., Abernathy, T. V., & Wiest, L. R. (Summer 2001). The reliability of students' ratings on faculty teaching effectiveness. *College Teaching* (49)3, pp 100-105.

Radmacher, S. A. & Martin, D. J. (May 2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *Journal of Psychology* (135)3, pp. 259-268.

Sax, L. J., Gilmartin, S. K., & Bryant, A. N. (August 2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education* (44)4, pp. 409 – 432.

Waters, M., Kemp, E., & Pucci, A. (December 1988). High and low faculty evaluations: Descriptions by students. *Teaching of Psychology* (15)4, pp. 203-204.



This work is licensed under a
[Creative Commons Attribution-NonCommercial-ShareAlike 2.5 License](http://creativecommons.org/licenses/by-nc-sa/2.5/).