

Students' Views of Course Evaluation

Yao-Yi Fu, Indiana University, Indianapolis
Suosheng Wang, Indiana University, Indianapolis
Ching-Hsu Huang, NPUST, Taiwan

Yao-Yi Fu and Suosheng Wang are Associate Professors of the Department of Tourism, Conventions, and Event Management at Indiana University, Indianapolis. Ching-Hsu Huang is Professor of the Department of Hotel and Restaurant Management at National Pingtung University of Science and Technology (NPUST), Taiwan.

Abstract

Students' evaluations of teaching have been widely used by educational institutions not only for ensuring quality of teaching, but also for making personnel decisions regarding faculty's tenure and promotion. There have been many concerns with the uses of the evaluations due to many validity and reliability problems with students' ratings of teaching. The purpose of this study is to understand how students view those issues. A survey of students' perceptions of course evaluations was conducted. The results of this research provide several insights about factors that can bias how students rate teaching performance.

Background

Students' evaluations of teaching have been widely used by colleges and universities throughout the world. Originally, teaching evaluations were developed as a way to assess and improve quality of teaching. Later, they were used for personnel decisions regarding faculty's teaching awards, merit pay, and tenure and promotion. More recently, they have been used as a quality assurance method that helps educational institutions to assess and demonstrate existence of effective teaching and student learning (Spooren, Brockx, & Mortelmans, 2013).

Although students' evaluations of teaching have been commonly utilized, these multiple uses of teaching evaluations have caused several concerns and controversies due to the validity and reliability of the evaluations. For instance, what exactly is evaluated by students has been questioned (Madden, Dillon, & Leak, 2010; Spooren, Brockx, & Mortelmans, 2013). Some scholars have pointed out that people have different interpretations of what constitutes effective teaching. Some evaluations may inadequately reflect valid measurement of teaching performance (Madden, Dillon, & Leak, 2010; Spooren, Brockx, & Mortelmans, 2013). Furthermore, it has been argued that students' ratings of teaching is more like a "customer satisfaction survey" which can be affected by many factors, such as students' expected grades, course difficulty and workload, students' interest in the subject, and instructors' personal traits and appearance (Knapper, 2001; Spooren, Brockx, & Mortelmans, 2013).

Purpose of the Study

A great amount of effort has been devoted to the research of college teaching evaluations. These studies have examined the reliability and validity of students' ratings of teaching and various factors that can bias the evaluations (Spooren, Brockx, & Mortelmans, 2013). Since the teaching evaluations are completed by students, it is important to look at students' perceptions of the evaluations because their views can provide many insights regarding concerns that have been raised by educators. Although several studies (Chen & Hoshower, 2003; Marlin, 1987) started looking at students' attitudes toward teaching evaluations, they focused mostly on students' motivation for participating in teaching evaluations and students' opinions on the administration of evaluations. This study attempts to gather students' perceptions related to more wide-ranging issues such as administration and use of course evaluations and the influence of potential biasing factors, like characteristics of instructors, on how students rate on the evaluations.

Methodology

A tourism department of a university in the United States conducted a survey online for this research. A total of 325 students enrolled in five undergraduate courses were asked to participate in the study. Those courses ranged from freshman to senior levels. Students enrolled in those five courses received a web link for the online survey. There were 198 students completed the survey, resulting in a response rate of 60.92%.

The survey items were developed based on concerns raised by previous research (Civian, & Brennan, 1996; Feldman, 1992; Ramsden, 1991; Ryan, Anderson, & Birchler, 1980; Wachtel, 1998). The survey items helped to collect students' opinions about administration, usage, and benefits of course evaluations, consequences of conducting teaching evaluations, and factors that can bias how students evaluate teaching. All survey items were measured on a five-point, Likert scale, ranging from "strongly disagree" to "strongly agree." In the literature, the terms 'course evaluations,' "teaching evaluations," and "student evaluations of teaching" were used interchangeably. On this study's survey instrument, "course evaluation" was adopted because it was the term that the tourism department used on the standardized teaching evaluation form.

Results

Findings of the survey are presented under different themes in this section. The means of all survey items are shown in Tables 1, 2, and 3 (next page)

Characteristics of the Student

The item that received the highest score in terms of research participants' agreement with survey items is "A student's motivation and interest prior to taking a course may impact the student's evaluation of the instructor." This finding supports previous studies (Feldman, 1978; Marsh & Cooper, 1981) that students with greater interest in the subject of a course prior to taking the course tend to give more favorable ratings on course evaluations.

The item that received the second highest mean rating is "A student's year of schooling (e.g. freshman, sophomore, junior, or senior) may impact the student's evaluation of the instructor." One can assume that students may evaluate teaching differently, as they move up to higher level courses, with increasing age, maturity, and knowledge. Under this category, what was ranked as the third highest is the item that concerns the impact of students' grades on course evaluations. With the mean score of 3.32 based on a five-point scale, the results indicate that a significant percentage of the participants in this study felt students' grades affected how they rated their instructors on the evaluations. An extensive amount of research (Badri, Abdulla, Kamali, & Dodeen, 2006; Blunt, 1991; Chacko, 1983; Nimmer & Stone, 1991; Spooren & Mortelmans, 2006; Tatro, 1995) has been devoted to examine the relationship between students' grades and ratings of instructors. Although conflicting results have been found, a significant number of studies have concluded that students who expected higher grades gave more favorable evaluations. Similarly, the result of this particular study implies that a considerable percentage of the survey participants admitted that grades did influence how they rated their faculty.

Table 1 Means of Items for Students' Perceptions toward Course Evaluation Characteristics of the Student and the Instructor. N = 198

Category	Survey Item	Mean
Characteristics of the Student (Average Mean: 3.19)	A student's motivation and interest prior to taking a course may impact the student's evaluation of the instructor.	3.63
	A student's year of schooling (e.g. freshman, sophomore, junior, or senior) may impact the student's evaluation of the instructor.	3.36
	Students' grades affect how they rate their instructors on course evaluations.	3.32
	A student's gender may impact the student's evaluation of the instructor.	2.44
Characteristics of the Instructor (Average Mean: 2.58)	Whether an instructor has a sense of humor or not affects how students rate him or her on course evaluations.	3.48
	An instructor's academic reputation may affect how students rate the instructor on the course evaluations.	3.43
	The amount of teaching experience of an instructor affects how students rate him or her on course evaluations.	2.65
	The physical appearance of an instructor affects how students rate him or her on course evaluations.	2.44
	Whether a lecturer is a full-time instructor or part-time instructor affects how students rate him or her on course evaluations.	2.37
	The position (lecturer, assistant professor, associate professor, or full professor) of an instructor affects how students rate him or her on course evaluations.	2.35
	Whether an instructor holds an administration job, such as school dean or department chair, affects how students rate him or her on course evaluations.	2.35
	The degree (Bachelor, Master's, Ph.D.) of an instructor affects how students rate him or her on course evaluations.	2.17
	The gender of an instructor affects how students rate him or her on course evaluations.	1.94

Note. All items were measured on a five-point, Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (5).

Table 2 Means of Items for Students' Perceptions toward Course Evaluation Characteristics of the Course and Administration of Course Evaluations N = 198

Category	Survey Item	Mean
Characteristics of the Course (Average Mean: 3.02)	The degree of difficulty of the course may impact a student's evaluation of the instructor.	3.82
	The level of the course (e.g. 100 - 200 low level or 300 - 400 high level) may impact a student's evaluation of the instructor.	3.40
	The subject of a course (statistics, geography, tourism, etc.) affects how students rate the instructor on course evaluations.	3.11
	The availability of the school's textbook resources may impact a student's evaluation of the instructor.	2.92
	Whether a course is required or an elective affects how students rate the instructor on course evaluations.	2.86
	The school's instructional facilities may impact a student's evaluation of the instructor.	2.86
	The classroom environment (noise from outside, room temperature, etc.) affects how students rate the instructor on course evaluations.	2.81
	The number of students in a class affects how students rate the instructor on course evaluations.	2.73
	The meeting time of a course affects how students rate the instructor on course evaluations.	2.64
	Administration of Course Evaluations (Average Mean: 3.40)	The responses on a student's course evaluation may be affected by the presence of the instructor.
The responses on a student's course evaluation may be affected by the presence of other students.		3.39
Students may feel some anxiety when filling out the course evaluation form.		3.04

Note. All items were measured on a five-point, Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (5).

Table 3 Means of Items for Students' Perceptions toward Course Evaluation
 Negative Impacts of Course Evaluations and Use of Course Evaluations N = 198

Category	Survey Item	Mean
Negative Impacts of Course Evaluations (Average Mean: 3.18)	An instructor may receive different ratings and comments from students which may cause difficulty for the instructor to decide the best way to teach.	3.94
	Instructors who provide quality teaching do not necessarily receive good teaching evaluations.	3.43
	Course evaluation may decrease instructors' morale about teaching.	3.09
	Course evaluation may negatively affect rapport between students and an instructor.	2.97
	Course evaluation may cause tension between students and an instructor.	2.83
	Course evaluation may lower an instructor's standards and expectations of students.	2.83
Use of Course Evaluations (Average Mean: 3.69)	Students' course evaluations provide an opportunity for an instructor's self-evaluation.	4.17
	Students' course evaluations can help to promote communication between students and instructors about teaching.	4.01
	Students' course evaluations can reflect the effectiveness of teaching.	3.94
	Students' course evaluations can motivate instructors to improve the quality of teaching.	3.93
	Students' course evaluations can be used as one of the evaluations for an instructor's job promotion.	3.72
	Results of course evaluation can be used as one of the references for teaching awards.	3.63
	Results of course evaluation should be useful in helping students to select what course sections to take.	3.60
	Results of course evaluation can be used as one criterion for renewing an instructor's job contract.	3.49
	Results of course evaluation can be shown in university publications that are accessible to students.	3.44
	Students' course evaluations help to improve the motivation of students to learn.	2.89

Note. All items were measured on a five-point, Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (5).

Characteristics of the Instructor

Much research (Allen, 1995; Feldman, 1987; Leventhal, Abrami, & Perry, 1976; Marsh & Ware, 1982; Naftulin, Ware, & Donnelly, 1973; Perry, Abrami, Leventhal, & Check, 1979) has been conducted to investigate if characteristics of instructors affected how students rated them on course evaluations. The participants of this study were asked to rate nine characteristics of instructors. The nine characteristics are: academic reputation, rank, degree, amount of teaching experience, gender, physical appearance, teaching full-time or part-time, holding an administrative job, and having a sense of humor.

The item that received the highest mean score is “Whether an instructor has a sense of humor or not affects how students rate on him or her course evaluations.” Although not much research has examined instructors’ sense of humor closely, a similar effect was found in an early study conducted by Naftulin, Ware, and Donnelly (1973). They found that instructors who were entertaining or expressive received higher teaching ratings. Marsh and Ware (1982) furthered examined this effect and stated that when students were not given incentive to learn, instructors’ expressiveness was found to have a greater influence on teaching evaluations than course content coverage. The item that received the second highest mean score concerns instructors’ academic reputation. This finding supports a study conducted by Leventhal, Abrami, and Perry (1976). In their study, it was found that students who selected classes based on instructors’ reputation gave higher ratings on course evaluations.

The rest of the seven characteristics of instructors all received mean scores below 3.0 based on a five-point, Likert scale. These lower ratings imply that the participants did not feel instructors’ rank, degree, amount of teaching experience, gender, physical appearance, teaching full-time or part-time, or holding an administration job had a significant influence on how students evaluated teaching.

Characteristics of the Course

Nine characteristics of courses were examined in this study, which are: course subject, difficulty of a course, number of students, whether a course is required or not, class meeting time, classroom environment, course level, availability of textbook resources, and instructional facilities. Three of these characteristics received mean scores above 3.0.

The item that received the highest rating is “The degree of difficulty of the course may impact a student’s evaluation of the instructor.” This finding confirms previous research (Dudley & Shawver, 1991) findings that there is a significant relationship between course difficulty and evaluations of teaching.

The item that received the second highest mean score is about the level of courses. Several previous studies (Feldman, 1978; Wachtel, 1998) found that higher level courses received higher ratings. Wachtel (1998) mentioned that several factors, such as age and maturity of students, class size, and electivity, may influence the connection between course level and teaching evaluations.

The third item that received a mean score above 3.0 is the subject of a course. Many previous studies (Cashin & Clegg, 1987; Feldman, 1978; Perry, Abrami, Leventhal, & Check, 1979; Ramsden, 1991) found a significant correlation between subject area and teaching evaluations. For instance, some researchers (Cashin & Clegg, 1987; Feldman, 1978; Perry, Abrami, Leventhal, & Check, 1979; Ramsden, 1991) found that courses such as mathematics and sciences that require quantitative skills received the lowest evaluations. However, this phenomenon needs to be further explored to see if there is a confounding effect from other factors, such as work load and grades.

Administration of Course Evaluations

The item that the survey participants most agreed with is “The responses on a student’s course evaluation may be affected by the presence of the instructor.” Several studies (Braskamp & Ory, 1994; Feldman, 1978) concluded that faculty’s presence in the classroom during administration of course evaluations was an important issue for students. That finding is echoed by the results of this study. Previous researchers (Braskamp & Ory, 1994) recommended that teaching evaluations should be distributed and collected by a third party in order to protect students’ anonymity.

The item that received the second highest mean score is “The responses on a student’s course evaluation may be affected by the presence of other students.” In addition, the

third item that received above a 3.0 mean rating is “Students may feel some anxiety when filling out the course evaluation form.” On the survey, there was a separate question that asked the survey respondents to indicate the best place for students to complete course evaluations. A total of 56.1% of them indicated that completion on the Internet is the best while 19.4% of them preferred to evaluate in a classroom. All these results imply that doing course evaluations alone on the Internet is the method that they felt most comfortable with. Moreover, some study participants commented that because course evaluations were usually given at the end of class, some students rushed through the evaluation just to leave as soon as possible without taking the evaluations seriously. One respondent commented that after doing course evaluations both in class and online, he or she felt online evaluations provided more honest answers.

Furthermore, there was an open-ended question in the survey that asked participants to provide any comments on course evaluations. This question revealed a few major suggestions that were recurring in the 93 comments that were collected. The most repeated one is that current evaluation questions are too general, which was mentioned by 23 survey participants. For instance, one comment stated that because the current evaluation form is the same for all courses, students may have just gone through the evaluation and gave answers without reading and thinking about the questions. One respondent pointed out that if the evaluation varies, students would have to read and think more carefully. Many participants recommend that evaluation questions need to be more specific toward different classes or at least towards different academic departments on campus. There were nine respondents who also suggested that more open-ended questions should be added.

Use of Course Evaluations

The study participants were asked to think about ten uses of conducting course evaluations. Items that received the two highest mean scores in terms of participants’ agreement were “Students’ course evaluations provide an opportunity for an instructor’s self evaluation” and “Students’ course evaluations can help to promote communication between students and instructors about teaching.” In addition, results of several other items indicate that students agreed that outcomes of course evaluations can be used to determine faculty’s promotion, teaching awards, and job contracts. Study participants also agreed that the results of course evaluations can be shown to students, which help them to select instructors.

These results echo what Chen and Hoshower (2003) found in their study of students’ motivation for participating in teaching evaluations. Their findings suggest that improving teaching and course content is a more important use of the evaluations than other uses, such as determination of faculty’s tenure, promotion, and salary raises.

Negative Impacts of Course Evaluations

Several questions were asked to see what students thought about the negative impacts of course evaluations on faculty. The item that participants most agreed with is “An instructor may receive different ratings and comments from students which may cause difficulty for instructors to decide the best way to teach.” This result implies that students recognized the fact that various suggestions for teaching may present a challenge for improving teaching.

The item that received the second highest mean is “Instructors who provide quality teaching do not necessarily receive good teaching evaluations.” This interesting result implies that research participants knew that course evaluations may not truly reflect the quality of teaching due to biasing factors.

These findings echo Beran and Rokosh’s (2009) study of faculty’s attitudes about teaching evaluations. Among 357 instructors who participated in their survey, nearly two thirds of

them did not agree with course evaluation results. Beran and Rokosh (2009) identified several reasons for the high percentage of the faculty's mistrust of teaching evaluations. Although faculty agreed that students should evaluate instructors, at the same time, they believed students are not knowledgeable enough to provide valid appraisals. Next, given that teaching is multifaceted, not all faculty place the same weight on the same aspects of teaching. Therefore, faculty may not agree with a solitary assessment of teaching effectiveness. As a result, many faculty may not value course evaluations much and would not change their teaching based on the evaluations.

Furthermore, Lindahl and Unger (2010) pointed out in their research that faculty may find cruelty in teaching evaluations, which can lead to many negative consequences. Their study focused on students' cruel comments, such as "This professor is evil." "This course ruined my senior year." "This class is a joke." These are caused by the "students-as-consumers" mentality and the methods and conditions of how evaluations are administered. Lindahl and Unger (2010) found this phenomenon can cause faculty's feelings of demoralization and anger. Subsequently, some faculty distrust and disengage in use of students' ratings of their teaching.

Conclusions

Students' evaluations of teaching have become routine practices in higher education. They are used for many purposes that can have tremendous impacts on colleges' and universities' teaching quality assurance and on faculty's tenure, promotion, and merit pay. Based on an anonymous survey conducted for this research, students who participated in the survey provided many candid opinions that indicate the evaluations can be problematic. For instance, one survey participant mentioned that course evaluations seem tilted in favor of students. The majority of students he or she was in class with will give a bad evaluation due to their lack of interest in a class or their inability to understand the instructor or course contents.

As discovered in this study, how students evaluate a course can not only be influenced by many factors, but the survey participants also admitted that they did not always take time to read the questions on teaching evaluations nor made thoughtful judgments.

In addition, there are many validity and reliability issues in the design of the evaluation instrument as revealed in several studies (Madden, Dillon, & Leak, 2010; Onwuegbuzie, Witcher, Collins, Filer, Wiedmaier, & Moore, 2007; Spooren, Brockx, & Mortelmans, 2013). Consequently, previous research (Beran & Rokosh, 2009; Nasser & Fresko, 2001; Spencer & Flyr, 1992) found that many instructors felt course evaluations did not reflect what constituted good teaching and provided little value for improvement of their teaching.

Although some of the issues discussed above may be resolved by adjusting the administration of the evaluations to give students a more suitable setting and more time to do course evaluations, many issues still require a great deal of further research and soul-searching discussions among students, faculty, and administrators about what constitutes good teaching and how it can be authentically assessed.

References

- Allen, M. (1995). Research productivity and positive teaching evaluations: Examining the relationship using meta-analysis. Paper presented at the Annual Meeting of the Western States Communication Association, Portland, Oregon.
- Badri, M. A., Abdulla, M., Kamali, M. A., & Dodeen, H. (2006). Identifying potential biasing variable in student evaluation of teaching in a newly accredited business program in the UAE. *The International Journal of Educational Management*, 20(1), 43-59.
- Beran, T. N., & Rokosh, J. L. (2009). Instructors' perspectives on the utility of student ratings of instruction. *Instructional Science*, 37, 171-184.
- Blunt, A. (1991). The effects of anonymity and manipulated grades on student ratings of instructors. *Community College Review*, 18, 48-54.

- Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work*. San Francisco: Jossey-Bass.
- Cashin, W. E., & Clegg, V. L. (1987). Are student ratings of different academic fields different? Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.
- Chacko, T. I. (1983). Student ratings of instruction: A function of grading standards. *Educational Research Quarterly*, 8, 19-25.
- Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education*, 28(1), 71-88.
- Civian, J. T., & Brennan, R. T. (1996). Student and course factors predicting satisfaction in undergraduate courses at Harvard University. Cerek Bok Center for Teaching and Learning, Harvard University. (ERIC Document Reproduction Service No. ED 394 440).
- Dudley, S., & Shawver, D. L. (1991). The effect of homework on students' perceptions of teaching effectiveness. *Journal of Education for Business*, 67, 21-25.
- Feldman, K. A. (1978). Course characteristics and college students' ratings of their teachers: What we know and what we don't. *Research in Higher Education*, 9, 199-242.
- Feldman, K. A. (1987). Research productivity and scholarly accomplishments: A review and exploration. *Research in Higher Education*, 26, 227-298.
- Feldman, K. A. (1992). College students' views of male and female college teachers: Part I—Evidence from the social laboratory and experiments. *Research in Higher Education*, 33, 317-375.
- Knapper, C. (2001). Broadening our approach to teaching evaluation. *New Directions for Teaching and Learning*, 88, 3-9.
- Leventhal, L., Abrami, P. C., & Perry, R. P. (1976). Do teacher rating forms reveal as much about students as about teachers? *Journal of Educational Psychology*, 68, 441-445.
- Lindahl, M. W., & Unger, M.L. (2010). Cruelty in student teaching evaluations. *College Teaching*, 58, 71-76.
- Madden, T. J., Dillon, W. R., & Leak, R. L. (2010). Students' evaluation of teaching: Concerns of item diagnosticity. *Journal of Marketing Education*, 32(3), 264-274.
- Marlin, J. (1987). Student perceptions of end-of-course evaluations. *The Journal of Higher Education*, 58(6), 704-716.
- Marsh, H. W., & Cooper, T. L. (1981). Prior subject interest, students' evaluations, and instructor effectiveness. *Multivariate Behavioral Research*, 16, 82-104.
- Marsh, H. W., & Ware, J. E. (1982). Effects of expressiveness, content coverage, and incentive on multidimensional student rating scales: New interpretations of the 'Dr. Fox' effect. *Journal of Educational Psychology*, 74, 126-134.
- Naftulin, D. H., Ware, J. E., & Donnelly, F. A. (1973). The Doctor Fox lecture: A paradigm of educational seduction. *Journal of Medical Education*, 48, 630-635.
- Nasser, F., & Fresko, B. (2001). Faculty views of student evaluation of college teaching. *Assessment & Evaluation in Higher Education*, 27, 187-198.
- Nimmer, J. G., & Stone, E. F. (1991). Effects of grading practices and time of rating on student ratings of faculty performance and student learning. *Research in Higher Education*, 32, 195-215.
- Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal*, 44(1), 113-160.
- Perry, R. P., Abrami, P. C., Leventhal, L., & Check, J. (1979). Instructor reputation: An expectancy relationship involving student ratings and achievement. *Journal of Educational Psychology*, 71, 776-787.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire. *Studies in Higher Education*, 16, 129-150.
- Ryan, J. J., Anderson, J. A., & Birchler, A. B. (1980). Student evaluation: The faculty responds. *Research in Higher Education*, 12, 317-333.
- Spencer, P. A., & Flyr, M. L. (1992). The formal evaluation as an impetus to classroom change: Myth or reality? (Research/Technical Report, Riverside, CA).
- Spooren, P. & Mortelmans, D. (2006). Teacher professionalism and student evaluation of teaching: Will better teachers receive higher ratings and will better students give higher ratings? *Educational Studies*, 32(2), 201-214.

- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83(4), 598-642.
- Tatro, C. N. (1995). Gender effects on student evaluations of faculty. *Journal of Research and Development in Education*, 28, 169-173.
- Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, 23(2), 191- 211.