What Influences End-of-Course Evaluations? Teaching & Learning versus Instrumental Factors

Abstract
Student evaluations of courses and teaching in the form of end-of-course surveys are ubiquitous in higher education, and at many institutions they serve as the primary basis for evaluating teaching effectiveness in the promotion and tenure process. Course evaluations of teaching are also controversial. It is often asserted that students use them to reward professors for easy courses and punish them for demanding ones, and many faculty believe that students’ evaluations are influenced by instrumental factors. This study investigates the relative influence of teaching and learning versus instrumental influences in students’ overall course evaluation ratings using data from a diverse sample of 44-year institutions.

Background & Purpose
A recent analysis found that at a diverse sample of 30 US institutions, 87-94% of respondents in different class levels indicated that they had submitted institution-provided course evaluations for their courses during the academic year (BrckaLorenz, McCormick, & Peck, 2014). Course evaluations of teaching are also controversial. It is often asserted that students use them to reward professors for easy courses and punish them for demanding ones (Boback & Marks, 2010), and many faculty believe that students’ evaluations are influenced by their expected grade. (Centra, 2003; Marsh, 1997; Mihaljević & Rustagi, 2008; Soja, Gupta, & Deter-Schmelz, 2002). Yet after multivariate analysis of data from more than 50,000 courses, Centra (2003) concluded, “teachers will not likely improve their evaluations from students by giving higher grades and less course work.” (p. 516).

Because course evaluations are widely used and consequential, it is important to gain deeper insight into how students formulate their ratings. Despite this question’s importance, the research base is surprisingly sparse, with a heavy focus on validity studies, most from the 1970s, 80s, and 90s (e.g., Marsh, Heiner, & Thomas, 1975; Marsh & Roche, 1997). A number of recent studies have investigated students’ perceptions about course evaluations (e.g., Mihaljević & Rustagi, 2008; Soja, Gupta, & Deter-Schmelz, 2002); these trends relate on single-institution or school convenience samples. The present study takes advantage of a multi-institutional survey in which students rated the impact of a range of 13 possible influences on their overall evaluations to investigate the relative influence of factors related to teaching and learning (such as the amount learned) versus instrumental factors (such as their expected grades).

Based on prior research (Centra, 2003; Soja, Gupta, and Deter-Schmelz, 2002), we hypothesize that students with evidence of a stronger orientation toward teaching and learning factors than instrumental ones. Because professional majors (e.g., business, engineering, health professions) are more closely tied to grades for which college grades figure large in post-college employment prospects, we also hypothesize that students in those fields will be more strongly oriented to instrumental factors than their counterparts in traditional liberal arts fields.

Data
A short set of questions about course evaluations was appended to NSSE at 44 U.S. institutions. Several questions probed views and behaviors related to instrument-provided course evaluations among the 88% of first-year and 95% of senior respondents who indicated having submitted evaluations for at least some of their courses during the academic year. One set of questions was included in the present study—students to rate how much various considerations influence their overall course evaluation ratings (Table 1).

Sample
The sample includes 4,408 (36%) first-year students and 7,890 (64%) seniors from institutions selected for the course evaluations set. Two-thirds of respondents were female (FY: 56%/SR: 63%), and nearly all first-year (87%) and half of seniors (55%) were of traditional college age. About half (52%) of first-year and three in five seniors (59%) were white and very few were Asian, Native Hawaiian, or Other Pacific Islander (FY: 9%/SR: 6%). Nearly all first-year students (89%) were enrolled full-time, over half (FY: 51%/SR: 57%) were first-generation college students (neither parent holds a bachelor’s degree), and nearly two-thirds (60%) of first-year and one in ten (18%) were living on campus. More than half (56%) of seniors had begun college elsewhere. Students in the sample were from a variety of majors with the largest proportions in health Professions (FY: 17%, SR: 13%); Business (FY: 12%/SR: 17%); and Biological, Agricultural, & Natural Resources (FY: 12%, SR: 9%). See Table 2 for additional details.

The sample represented considerable institutional diversity. Almost three-quarters (72%) of sample members attended public institutions; about one-quarter (28%) attended private non-profit institutions; about one in ten (11%) went to Catholic schools; about one in five (20%) went to Master’s granting colleges and universities, and a comparable share (38%) at doctorate-granting universities. Another one-fifth attended baccalaureate colleges (39%), with the remainder (5%) unclassified or special focus institutions. One in five students (21%) were at institutions enrolling fewer than 2,500 total undergraduate enrollment and another one-fifth (20%) were at medium sized institutions (2,504-6,099). Two-fifths (41%) were at large institutions (5,000-9,999), and another one-fifth (16%) attended very large institutions (10,000 or more).

Measures
The dependent variables for this study were two scales derived from the 13 items in Table 1 using exploratory factor analysis with oblique rotation: Teaching & Learning Factors (TLF) (α = .978) and Instrumental Factors (IF) (α = .947).

Research Questions & Results
1. What is the relative influence of teaching and learning versus instrumental influences in students’ overall course evaluation ratings?

We found that TLF and IF are positively correlated (r = .48, p < .001). First-year students average TLF and IF scores were 48.6 and 36.9, respectively. Fifty percent of .12 and 14.5, respectively. The TLF was about half again as large for seniors who had an average TLF score of 47.59 versus IF scores of 3.74 (SO: 12.70 & 15.22).

Thus for both groups, teaching & learning factors exert a stronger influence on course evaluation ratings than do instrumental considerations, and the difference is stronger for seniors.

2. What student characteristics predict teaching and learning influences on students’ course evaluation ratings?

A host of student characteristics are associated with increased influence on overall course evaluation ratings by Teaching & Learning Factors, holding other student and instrumental characteristics constant (Table 3). These students are slightly more influenced by TLF (β = .034, p = .012). Spending more time preparing for class (β = .556, p < .001) and increased subjective levels of course challenge (β = .717, p < .001) are positively related to increased influence by TLF. Liberal arts majors are more influenced by TLF (β = -.717, p < .001). Pronounced differences were found for gender and class level, with both female students (β = 1.068, p < .001) and seniors showing higher levels of TLF (β = 2.795, p < .001).

The largest negative differences are related to ethnic/racial identification. Asian, Native Hawaiian, or Other Pacific Islander (β = -2.795, p < .001) and Black or African American (β = -5.275, p < .001) were more strongly influenced by IF. First-generation students were more influenced by instrumental considerations (β = -0.696, p = .009). Interestingly, course challenge was also positively related to IF (β = .1, p < .001). Again, the largest differences were related to racial/ethnic identification. Asian, Native Hawaiian, or Other Pacific Islander (β = -3.359, p < .001); Hispanic or Latino (β = -2.364, p < .001); and American Indian, Alaskan Native, Other, or Multiracial (β = -5.935, p < .001) were all more influenced by IF compared to their respective White peers. There is no evidence of gender difference for IF, and the effects of institutional factors were not significant or were small.

3. What student characteristics predict instrumental factors on students’ course evaluation ratings?

A similarly wide variety of student characteristics also predicted increased influence on overall course evaluation ratings Instrumental Factors. Older students (β = -.189, p < .001) and students who were single (β = -.324, p < .001) were less influenced by IF. First-generation students were more influenced by instrumental considerations (β = -0.696, p = .009). Interestingly, course challenge was also positively related to IF (β = .1, p < .001). Again, the largest differences were related to racial/ethnic identification. Asian, Native Hawaiian, or Other Pacific Islander (β = -3.359, p < .001); Hispanic or Latino (β = -2.75, p < .001); Black or African American (β = 1.359, p < .001) and American Indian, Alaskan Native, Other, or Multiracial students (β = 1.955, p < .001) were all more influenced by IF compared to their White peers. There is no evidence of gender difference for IF, and the effects of institutional factors were not significant or were small.

Discussion & Significance
These results offer both reassurance and caution. On the reassuring side, our findings support the contention that students pay more heed to teaching and learning factors than to instrumental considerations when assigning overall course ratings, and this is especially so for women, older students, and more advanced students. Students who take their education seriously, as manifested by study time, and students who feel their courses challenge them to seriously, as manifested by study time, and students who feel their courses challenge them to

Appendix A

Table 1: NSSE 2013 Course Evaluation Experimental Items (excerpt)

Table 2: OLS Regressions Predicting Teaching & Learning Influences on Overall Course Evaluation Ratings

Table 3: OLS Regressions Predicting Teaching & Learning and Instrumental Influences on Students’ Course Evaluation Ratings

Discussion & Significance
These results offer both reassurance and caution. On the reassuring side, our findings support the contention that students pay more heed to teaching and learning factors than to instrumental considerations when assigning overall course ratings, and this is especially so for women, older students, and more advanced students. Students who take their education seriously, as manifested by study time, and students who feel their courses challenge them to

See more about NSSE and this study at nsse.iub.edu/html/publications_presentations.cfm