

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 their expected grade. 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 4-year institutions.

Background & Purpose

A recent analysis found that at a diverse sample of 30 US 4-year institutions, 87-94% of respondents in different class levels indicated that they had submitted institution-provided course evaluations for at least some of 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 (Babcock & Marks, 2010), and many faculty believe that students' evaluations are influenced by their expected grade (Centra, 2003; Marsh, 1987; Mukherji & Rustagi, 2008; Sojka, Gupta, & Deeter-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, Fleiner, & Thomas, 1975; Marsh & Roche, 1997). A number of recent studies have investigated student and faculty perceptions about course evaluations (e.g., Mukherji & Rustagi, 2008; Sojka, Gupta, & Deeter-Schmelz, 2002) but these tend to rely 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 grade).

Based on prior research (Centra, 2003; Sojka, Gupta, and Deeter-Schmelz, 2002), we hypothesize that students will evidence a stronger orientation to teaching and learning factors than instrumental ones. Because professional majors (e.g., business, engineering, health professions) are more closely tied to the labor market and for which college grades figure in post-collegiate 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 institution-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—the focus of the present study—asked 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: 63%/SR: 63%), and nearly all first-years (87%) and half of seniors (55%) were of traditional college age. About half (52%) of first-years and three in five seniors (59%) were white and very few were international students (FY: 9%/SR: 6%). Nearly all first years (95%) and three in four seniors (77%) were enrolled full time, over half (FY: 53%/SR: 57%) were first-generation college students (neither parent holds a bachelor's degree), and nearly two-thirds (60%) of first-years and one senior in ten (14%) 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 Sciences, Agriculture, & 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. Nearly two-fifths (38%) were at Master's-granting colleges and universities, and a comparable share (38%) at doctorate-granting universities. Another one-fifth attended baccalaureate colleges (19%), with the remainder (5%) at 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 (2,500-4,999). 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) ($\alpha = .978$) and Instrumental Factors (IF) ($\alpha = .947$).

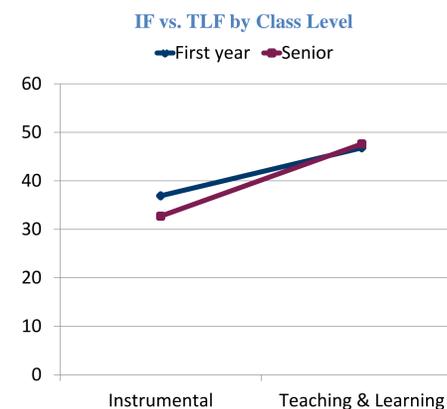
Table 1: NSSE 2013 Course Evaluation Experimental Items (excerpt)

	TLF	IF
How much do the following influence your overall ratings on end-of-course evaluations? <i>Very much, Quite a bit, Some, Very little</i>		
How much you learned	X	
The total amount of effort required		X
The grade you expect to receive		X
The ease or difficulty of assignments and exams		X
Whether the course was required or an elective		X
The feedback you received	X	
How much course sessions interested you		X
How much assignments (readings, homework, labs, etc.) interested you		X
The instructor's knowledge of course content	X	
The ease or difficulty of understanding the instructor	X	
How effectively the instructor used course time	X	
How clearly the instructor explained difficult material	X	
How the instructor interacted with students	X	

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 46.81 and 36.95, respectively (SD of 12.66 & 14.54). The gap was about half again as large for seniors who had an average TLF score of 47.59 versus average IF of 32.74 (SD 12.27 & 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 institutional characteristics constant (Table 3). Older students are slightly more influenced by TLF ($B = .034$, $p = .012$). Spending more time preparing for class ($B = .556$, $p < .001$) and increasing subjective levels of course challenge ($B = .701$, $p < .001$) are positively related to increased influence by TLF. Liberal arts majors are more influenced by TLF ($B = .717$, $p = .001$). Pronounced differences were found for gender and class level, with both female students ($B = 1.069$, $p < .001$) and seniors showing higher levels of TLF ($B = 1.805$, $p < .001$).

The largest negative differences are related to racial/ethnic identification. Asian, Native Hawaiian, or Other Pacific Islander ($B = -5.240$, $p < .001$); Black or African American ($B = -3.559$, $p < .001$); Hispanic or Latino ($B = -2.364$, $p < .001$); and American Indian, Alaska Native, Other, or Multiracial ($B = -.758$, $p = .043$) students were less influenced by teaching and learning factors than their White peers, and some of these effects were quite large. No differences were found by first-generation status, nor did institutional characteristics contribute to the importance of TLF.

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 ($B = -.189$, $p < .001$) and students who spend more time preparing for class ($B = -.525$, $p < .001$) were less influenced by IF components. Seniors ($B = -2.795$, $p < .001$) and liberal arts majors ($B = -1.574$, $p < .001$) were notably less influenced by IF.

First-generation students were more influenced by instrumental considerations ($B = .696$, $p = .009$). Interestingly, course challenge was also positively related to IF ($B = 1.123$, $p < .001$). Again, the largest differences were related to racial/ethnic identification. Asian, Native Hawaiian, or Other Pacific Islander ($B = 8.342$, $p < .001$); Hispanic or Latino ($B = 5.725$, $p < .001$); Black or African American ($B = 5.270$, $p < .001$); and American Indian, Alaska Native, Other, or Multiracial students ($B = 1.955$, $p < .001$) were all more influenced by IF components than their White peers. There is no evidence of gender difference for IF, and the effects of institutional characteristics were not statistically significant or were trivial.

Table 3. OLS Regressions Predicting Teaching & Learning and Instrumental Influences on Overall Course Evaluation Ratings

	Teaching & Learning Factors			Instrumental Factors			
	Unst. B	Std. Error	Sig.	Unst. B	Std. Error	Sig.	
(Constant)	25.177	.776	.000	6.109	.981	.000	
Instrumental influences score	.408	.007	.000	--	--	--	
Teaching & learning influences score	--	--	--	.591	.011	.000	
Age	.034	.013	.012	-.189	.016	.000	
Female	1.069	.226	.000	-.281	.273	.303	
First-generation	-.102	.222	.645	.696	.267	.009	
Senior	1.805	.252	.000	-2.795	.302	.000	
Liberal Arts Major	.717	.224	.001	-1.574	.269	.000	
	Asian, Native Hawaiian, or Other Pacific Islander	-5.240	.468	.000	8.342	.561	.000
	Black or African American	-3.559	.333	.000	5.270	.399	.000
	Hispanic or Latino	-2.364	.425	.000	5.725	.509	.000
	American Indian, Alaska Native, Other, Multiracial	-.758	.374	.043	1.955	.449	.000
	I prefer not to respond	-.232	.535	.664	.149	.644	.817
	Weekly time spent preparing for class	.556	.063	.000	-.525	.076	.000
	Course challenge (1=low, 7=high)	.701	.095	.000	1.123	.114	.000
Private	.072	.341	.834	.142	.410	.729	
	Master's Colleges and Universities	-.227	.306	.457	.669	.368	.069
	Baccalaureate Colleges	.350	.422	.407	.369	.507	.468
	Very Small (< 1,000)	-1.342	.746	.072	.414	.898	.644
	Small (1,000-2,499)	-1.055	.582	.070	1.409	.700	.044
	Medium (2,500-4,999)	-.460	.464	.322	.190	.559	.734
	Large (5,000-9,999)	.391	.368	.287	.497	.442	.262
	Competitive	.443	.352	.209	.261	.424	.538
	Most/Highly Competitive	.162	.387	.676	-.604	.466	.195
F	168.05 (p<.001)			196.33 (p<.001)			
Adjusted R-square	.273			.305			

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 do their best also evidence a stronger influence of teaching and learning factors. On the concerning side, however, we find that students in professional majors are less influenced by teaching and learning considerations and more influenced by instrumental ones, relative to peers in liberal arts majors. This is a worrisome finding given the well documented growth of professional majors.

Most concerning, however, are the findings that nonwhite students evidence considerably lesser influence of teaching & learning factors and greater influence of instrumental considerations. These worrisome findings are deserving of further investigation to understand the apparent strong differences in orientation to the important task of evaluating courses.