

The School of Education's (SOE) fall 2012 implementation of LMU's first large-scale online teaching evaluation system provides an opportunity to answer questions about possible effects of transitioning from paper to online collection. Decision makers considering such a change may be faced with concerns about online evaluations having different (1) mean scores, (2) response rates, and (3) types of students likely to respond (e.g., the happy or upset). This brief shares results of a study conducted by Institutional Research (IR) to shed light on these questions.

Method

To explore these questions, we constructed three outcomes at the course section level. First, we calculated μ —the mean overall instructor effectiveness rating. Another IR study found this global measure to be correlated with each of the other items between 0.78 and 0.90. Second, we constructed σ , the standard deviation of the global item. This measured whether responses became more extreme. Third, we calculated response rate, R , by dividing the number of ballots received by end-of-term enrollment. Response rate was capped at 100% and we eliminated from the data two sections that had a 0% rate despite verifying that ballots were distributed.

We estimated the effect of online collection by comparing outcomes before and after implementation. However, this comparison was made only within course-instructor *lines*. A line was composed of all same-course sections taught by a single instructor that were evaluated in fall 2012 and at least once between fall 2009 and spring 2012—the period for which the most-recent teaching evaluation form was used. Courses with the same number, subject code, catalog title, and instructor code were put in the same line, but to be comprehensive the SOE confirmed when courses with changes in numbering, subject code, or catalog title should also be included. These lines constituted the basis of the before/after comparison, making it unlikely that any results could be due to differences in instructor or curriculum across the years.

The final data set contained 74 lines with 2 to 15 sections each (Appendix A) for a total of 308 sections. It represented 56 unique courses taught by 67 unique instructors. With the design of this study aimed at maximizing internal validity, its findings might not generalize to the entire set of SOE courses; of 193 sections evaluated in fall 2012, just 87 sections were included here. The reasons for this smaller set were that online courses were not evaluated before fall 2012; team-taught courses tended not to have the same instructors in the past; and new courses, instructors, and course preparations for returning instructors did not have historical comparisons. Means and ranges of all variables were listed in Appendix B.

The statistical model used to estimate the effect of online collection was a fixed-effect regression, clustered on instructor, with analytic weights to account for variation in the number of ballots received (for the μ and σ models) and enrollment (for the R model). In a previous study, IR ascertained that average student interest was not affected by the new online collection system, but was a powerful predictor of overall instructor effectiveness ratings. Thus, we used it as a control variable to

account for natural variation in student interest from fall to spring, as well as random variation within line. Online collection was coded binary—“yes” if and only if the section was taught in fall 2012. The model was specified as follows.

$$\begin{Bmatrix} \mu_i \\ \sigma_i \\ R_i \end{Bmatrix} = \beta_0 + \beta_1 \text{Online}_i + \beta_2 \overline{\text{Interest}}_i + \sum_{j=2}^{74} \beta_{3,j} \text{Line}_{i,j} + \varepsilon_i$$

Results

Model results varied by outcome and were placed in the following table. Only response rate was clearly impacted, with an average decline of 11 percentage points ($p < 0.01$). For context, the average response rates for sections evaluated by paper and online were roughly 93% and 82%, respectively. The standard deviation of overall instructor effectiveness was not shown to have a relationship to online collection. The data did not suggest that respondents were more varied in their ratings. Finally, the mean score for overall instructor effectiveness increased by 0.09 points on average, but fell just outside of statistical significance ($p < 0.07$).

	μ		σ		R	
Variable	β	p	B	p	β	p
Online	0.09	0.07	-0.06	0.16	-0.11	0.00
Interest	0.65	0.00	-0.50	0.00	0.02	0.59
R ²	68.6%		56.5%		46.9%	

Discussion

Central administrators and others in the SOE implemented an online course evaluation system in fall 2012 and at this time there is not enough evidence to suggest that it affected instructors’ average rating or distribution of ratings. This result comes despite a lower response rate, suggesting that sacrificing even a fair chunk of ratings data on what must be considered an extremely high 93% base response rate can still yield information that is both representative of the population and consistent with the past. Recognizing that students would no longer be held in the classroom to fill out paper ballots and might choose not to respond, those involved attempted to bolster response rates in many ways, including email alerts, space in computer labs dedicated for evaluations, and encouragement of instructors to relay the importance of evaluations to students. Administrators moving their evaluation system online should be conscious of past response rates and how they will ensure an adequate number of replies, but appear to be free to move forward without worrying about whether it will fundamentally change ratings.

Appendix A – Sections Included in Data Set

Course/Instructor “Line”		Online	Paper Evaluations					
2012 Course	Instructor	F12	S12	F11	S11	F10	S10	F09
EDES 401	I. Higareda	1	1	1				
EDES 401	M. Reyes	1		1	1			1
EDES 434	L. Knatcal	1		1	1	1		1
EDES 5001	I. Higareda	2	2	2	2	2		2
EDES 5003	C. Fried	1				1		
EDES 5003	J. Green	2		1	1	2	2	2
EDES 5253	K. Berlin	1		1		1		
EDES 5254	D. Rodriguez	1		1		1		
EDES 5254	M. Ryan	1		1				
EDES 5255	A. Boyle	1		1		1		
EDES 5255	L. DeVeney	1				1		
EDES 5400	L. Ponciano	1		1		1		1
EDES 5401	A. Shabazian	1		1		1		1
EDES 5406	L. Ponciano	1		1		1		1
EDES 5407	L. Massa	1		1				
EDES 6103	A. Serrano	1	1		1	1	1	1
EDES 6200	L. Knatcal	1	2	1	2	1	2	1
EDES 6200	C. Watson	1	1	1	1	1	1	
EDES 6202	C. Watson	1		1				
EDES 6203	S. Gacina	1	1	1		1		
EDLA 489	D. Faill	1		1		1		
EDLA 489	M. Mennell	1	1	1	1	1		
EDLA 489	A. Niemi	1				1		
EDLA 491	S. Taylor	1	1	1	1	1		
EDLA 6421	E. Pack	1		2		2	1	2
EDLA 6422	S. Ali	1		1				
EDLA 6422	J. Elliot	1		1		1		1
EDLA 6422	A. Sabatino	1				1		
EDLA 6425	J. Isken	1		1		1	1	
EDLA 6427	L. Busia	1	1			1	1	
EDLA 6428	N. Orlando	1		1				
EDLA 7021	K. Huchting	1		1		1		
EDLA 7049	M. Lavadenz	1				1		
EDLA 7049	E. Reilly	1		1				
EDSS 440	M. Captain	1		1			1	1
EDSS 6001	J. Gomez	2		1				
EDSS 6001	A. Herrick	1		1				
EDSS 6300	J. Tibbetts	1		1				1
EDSS 6302	T. Aceves	1				1		
EDSS 6302	K. Doyon	2		1				
EDSS 6325	M. Azzi	1		1				
EDSS 6325	K. Goldman	1		1				

2012 Course	Instructor	F12	S12	F11	S11	F10	S10	F09
EDSS 6362	P. DeSena	1	1	1		1		1
EDSS 6368	E. Mejia	1	1	1	1	1		1
EDSS 6376	W. Welcher	1	1	1				
EDSS 6379	W. Parham	1	1	1				
EDSS 6382	K. Hawkins	1	1	1	1	1		
EDSS 6386	P. DeSena	1	1	1	1	1	1	1
EDSS 6386	C. Haddy	1	2	1		1	1	1
EDSS 6390	M. Jung	2	1	1				
EDSS 6393	S. Bailey	1	1	1				
EDSS 6394	D. Hehr	1	1	1				
EDSS 6500	L. Pikover	1		1				
EDSS 6504	E. Garcia	1		1				
EDSS 6506	R. Grant	1		1		1		1
EDSS 6508	E. Holman	1		1		1		
EDSS 6520	K. Hawkins	1				1		1
EDSS 6524	J. Pollick	2		2		2		2
EDSS 6528	S. Mora	1		1		1		
EDSS 6538	B. Leung	1				1		1
EDUR 400	M. Schauko...	1				1		
EDUR 5000	E. Colin	2	2	2	3	2	2	2
EDUR 5000	M. Sanchez	2	2	2				
EDUR 5004	L. Costello	1		1		1		
EDUR 5004	W. Crean	2		2		1		1
EDUR 5004	M. Thomas	1		1		1		
EDUR 5014	R. Patrick	2		2		1		
EDUR 5014	M. Ryan	3		2		2		
EDUR 5016	L. Rosenthal	2	2	2	2	1		
EDUR 5018	M. Gottfried	1		2				
EDUR 5501	K. Laskasky	1		1				
EDUR 6960	G. Knotts	1						1
EDUR 6961	E. Brewer	1	2	1	1			1
EDUR 6968	M. Perez	1	1	1	1	1		

Appendix B – Summary of Variables

Variable (At the Section Level)	Mean	S.D.	Min.	Max.
Overall Instructor Effectiveness (μ)	4.54	0.46	2.56	5.00
Overall Instructor Effectiveness (σ)	0.55	0.37	0.00	1.56
Response Rate (R)	0.90	0.13	0.40	1.00
Online Evaluation (Yes or No)	0.28	0.45	0	1
Average Student Interest	4.36	0.41	2.60	5.00
Overall Inst. Effectiveness Item Responses	13.7	5.72	2	31
Enrollment in Section	15.4	6.13	3	35