OIRSA | LEHIGH UNIVERSITY

Office of Institutional Research & Strategic Analytics

Successful Strategies to Encourage Student Participation in Course Evaluations

Date: November 16, 2022

Prepared by: Benjamin Pichlmaier, Research Analyst

Report summary: Course evaluations help improve instruction efforts and inform hiring and tenure decision. Low response rates, on the other hand, can induce response bias and thus impact the quality of evaluation data.

To help increase student response rates, OIRSA asked instructors after the end of the Spring term 2022 in an online survey what they did in order to encourage their students to participate in course evaluations. This report summarizes the results and details the analyses OIRSA performed on the survey data. The most important finding is that any action that instructors took to motivate students was helpful. Providing students extra time in-class to complete course evaluations proved to be especially effective.

1. Summary

After the Spring term 2022, OIRSA asked faculty at Lehigh University what, if anything, they did in order to motivate students to participate in course evaluations. 25% of 736 invited instructors completed the survey. In our analyses we compared mean response rates and employed stepwise hierarchical regressions in order to detect effects of instructors' actions.

While 16.5% of instructors reported that they did not do anything to encourage students to participate in course evaluations, our analyses revealed that utilizing any strategy to encourage participation resulted in a significant 22.5% response rate increase (36.4% vs. 58.9%, $p<.001$). The most impactful approach was providing students with time during class to complete the evaluations (resulting in an average increase from 46.2% to 67.0%), followed by reminding them to take the survey (increase by 8.5%), reiterating the use of course evaluations (plus 7.1%), or emphasizing confidentiality (plus 5.1%). We also found substantial increases for instructors who offered a collective reward (e.g., a small extra course credit or dropping homework) if the class met or exceeded a certain response rate threshold. However, this strategy was used in only ten cases which compromises a generalization of this finding.

Hierarchical stepwise regressions indicated that providing extra time was the best predictor for response rates. The best fitting model explained 19.4% of variation in response rates and predicted for a baseline response rate (intercept) of 41.8% an increase of 20.3% for providing extra time and 6.7% for reminding students to take the evaluation survey, irrespective of class size. Reiterating confidentiality and the use of evaluation results did not show substantive contributions in models with more than 2 predictors, which is most likely due to shared variance with "reminding students of course evaluations" (instructors who do so probably also discuss the use of evaluation results and confidentiality). The lack of statistical significance in these models therefore reflects rather methodological issues than the importance of these practices. Lastly, our sample and the entire invited population did not differ in response rates or course sizes, supporting the generalizability of these findings to Lehigh University.

2. Methods

The survey was administered as an online instrument from May 17th to May 31st 2022. 184 of 736 of all invited instructors completed the survey (25% response rate) and their responses represented 309 of 1,152 (26.8%) of all evaluated courses (courses that were cross listed or combined for other reasons counted as one course; early evaluated courses or practicums were not included). Participants were invited by email and asked to indicate what they did to encourage course evaluation completion for a

given course on a multiple-choice item with the following response selection: giving students in-class time to complete evaluations, reminding students that evaluations are open, reiterating how student feedback will be used, reminding students that their responses are confidential, the choice ”other” (with the request to share what activity was employed), and the option to indicate that one did not do anything in this regard. The question was repeated for each course an instructor had taught in the Spring 2022 term. For a very small number of instructors with a high teaching volume, we combined similar courses (e.g., same topic and format) in order to avoid a too burdensome survey length. These collapsed courses were disentangled during data cleaning.

The basic unit for analyses was the unique combination of course and instructor that participated in the survey (184 instructors covering 309 courses). We coded all open-ended answers to the write-in option “other” and included the derived categories in our quantitative analyses for all responses with 5 or more entries. This resulted in the following categories: offering collective course credits or other incentives, reminding students, and stressing the importance of course evaluations. All analyses were performed with R. We tested all applying assumptions to perform regression analyses and tests for mean differences. Multicollinearity was not of concern, but the distribution of response rates (the dependent variable) was skewed to the right, indicating that the majority of courses had a lower response rate. For that reason, we used the non-parametric Wilcoxon signed rank test to test mean differences. For regression analyses, we also inspected Q-Q plots for the best model resulting from hierarchical regressions, which did not result in concerns. We also analyzed alternatively selected models and the final model with robust non-parametric regression approaches (R package MASS), which did not produce substantially different results. Finally, we compared the mean enrollment number per course and response rates of the respondent sample to the entire population (all course-teacher combination for the Spring term 2022) by use of the Wilcoxon signed rank test and did not find any differences. The results from these non-parametric approaches give us confidence that the skewed response rate distribution did not distort our analyses.

3. Results

The mean response rate in the survey was 55.2% (Median=53.7%, *SD*=24.1). We compared the mean response rate for each group of instructors who used a specific approach with the complement group who did not use it. These differences were in all cases significant ($p<.001$) in favor of the use of the strategy, with differences ranging from 5.1% for emphasizing confidentiality to 20.8% for giving time to complete course evaluations during class (see table 1 for more details). In other words: doing anything was always better than doing nothing. Some instructors offered students collective incentives in the

Table 1

Mean Differences in Response Rates for different motivation strategies

	Count (Percent)	Mean	SD	% Diff
Gave extra time	133 (43.0%)	67.0	21.8	
Did not give extra time	176 (67.0%)	46.2	21.9	20.8
Reminded	214 (69.3%)	57.8	23.2	
Did not remind	95 (30.7%)	49.3	25.4	8.5
Reiterate use	164 (53.1%)	58.5	23.4	
Use not reiterated	145 (46.9%)	51.4	24.6	7.1
Stressed confidentiality	123 (38.8%)	58.3	24.5	
Confidentiality not stressed	186 (60.2%)	53.2	23.8	5.1
Did any of the above	258 (83.5%)	58.9	23.7	
Did not do any of the above	51 (16.5%)	36.4	16.7	22.5
Collective extra credit	10 (3.2%)	83.3	12.4	
No extra credit	299 (96.8%)	54.2	23.9	29.1
Any collective incentives	14 (4.5%)	84.0	12.1	
No incentive	295 (95.5%)	53.8	23.7	30.2

Note: All differences are statistically significant with $p<.001$

form of a small extra course credit or dropping homework for the entire class if response rates exceeded a certain threshold. We report the results of mean comparisons for this practice, but the small number of instructors reporting the use of incentives does not allow generalization to the entire population. Our findings are in line with higher education research (e.g., Sundstrom et al., 2016); however, the effects of this practice on the quality of course evaluations are unclear. Other practices that we did not test statistically due to the small number of practitioners included embedding course evaluations in the course work, requesting students to bring their computer to class (to participate in evaluations), and making participation in evaluations a course requirement.

All strategies to motivate student participation correlated positively with response rates, while the absence of any efforts was systematically associated with lower participation (see table 2 for more details). Enrollment numbers and response rates had a small significant inverse relation ($r(309) = -.18$, $p<.01$).

In order investigate effects of course size, we first tested the difference in response rates in our sample on the median split. Classes with enrollment above the median (n_1=154, SD_1=23.3) showed in a Wilcoxon rank sum test a smaller mean response rate (x_1=51.4, $p<.01$) than smaller classes (n_2=155, SD_2=23.3, x_2=58.9). The comparison of the average enrollment for instructors who reported not having done anything to encourage student participation (n_1=51) to those who did do so (n_2=258) showed also substantial differences ($p<.01$). No efforts corresponded to larger courses (median$_1$=25, SD=49.4) and vice versa (median$_2$=20, SD_2=18.5); this effect remained substantial once we tested a sample that eliminated outliers (class size > 100) due to the high standard deviation of "no effort" courses.

Table 2

Means, standard deviations, and correlations with confidence intervals

Variable	*M*	*SD*	1	2	3	4	5	6	7
1. Enrollment	26.56	26.43							
2. Response rates	55.16	24.14	-.18** [-.28, -.07]						
3. No action taken	0.17	0.37	.16** [.05, .26]	-.35** [-.44, -.24]					
4. Reminder	0.69	0.46	-.10 [-.20, .02]	.16** [.05, .27]	-.67** [-.72, -.60]				
5. Gave extra time	0.43	0.50	-.10 [-.20, .02]	.43** [.33, .51]	-.39** [-.48, -.29]	.08 [-.03, .19]			
6. Reiterated use	0.53	0.50	.02 [-.09, .13]	.15* [.03, .25]	-.47** [-.56, -.38]	.51** [.42, .59]	.08 [-.03, .19]		
7. Emphasized confidentiality	0.40	0.49	-.10 [-.21, .01]	.10 [-.01, .21]	-.36** [-.45, -.26]	.46** [.36, .54]	.19** [.08, .29]	.38** [.28, .47]	
8. Collective incentives	0.05	0.21	.02 [-.09, .13]	.26** [.15, .36]	-.10 [-.21, .01]	.15* [.03, .25]	-.00 [-.11, .11]	.17** [.06, .28]	.11 [-.00, .22]

Note. M and *SD* are used to represent mean and standard deviation, respectively. 3. to 8. Constitute dummy coded variables, means and standard deviations should therefore not be interpreted. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates $p < .05$. ** indicates $p < .01$.

In order to investigate if this can be interpreted as a tendency of instructors with large classes to invest less effort in motivating students to participate in course evaluations, we analyzed this relation with mediation and moderation hypothesis (our understanding and interpretation of causal inference from this kind of analyses follows Hayes, 2017). In the mediation analysis with activities to motivate students to participate in course evaluations as a mediator of effects from enrollment on response rates, we found that the mediator reduced the relationship between enrollment numbers and response rates considerably; however, enrollment numbers did not predict the mediator in this model. When we tested a moderation hypothesis for effects from course size and motivation activities on response rates, the interaction term and class size were not significant. Taken together, this suggests that an instructor's engagement in such motivational activities should be seen as a factor that influences response rates independently of the size of a class.

As a next step we performed hierarchical regressions (stepwise inclusion) in order to identify the contribution of each strategy to increases in response rates. We report here only the best fitting model. This model explained 19.4% of variation response rates; providing students extra time during class added 20.3% (p<.001), and reminding students to take the survey another 6.7% (p<.01) to a baseline (intercept) of 41.8%. Significance and size of these effects did not change substantially when controlled for course enrollment numbers (see also table 3). When reminding students to take the survey and giving extra time were entered in the regression equation simultaneously with reiterating the use of course evaluations and assuring confidentiality, the last two variables were no longer significant. However, this does not mean that these practices are not important; it is rather reasonable to assume that instructors who send reminders to students also emphasize the confidentiality and use of course evaluation data, which would then produce redundancy once the variables are considered together. Accordingly, we found in single regressions significant effects for stressing the use of evaluations (β=7.05, SE=2.73, p<.01) and addressing confidentiality concerns (β=5.14, SE=2.79, p<.05). Not doing anything, on the

Table 3

Final regression model (stepwise inclusion), controlling for enrollment per class

	Regression 1	Regression 2	Regression 3
Intercept	49.30 (2.45)***	41.81 (2.41)***	45.56 (2.82)***
Reminder	8.47 (2.94)**	6.67 (2.68)*	6.07 (2.67)*
Extra time		20.31 (2.50)***	19.76 (2.48)***
Enrollment			-0.13 (0.05)*
$R^2_{adjusted}$	0.03	0.19	0.21
R^2_{change}	0.03	0.16	0.02

Note. Standard errors are in parentheses.

* indicates p < .05. ** indicates p < .01 *** indicates p < .01

other hand, predicted clearly decreases in response rates (β=-22.53, *SE*=3.48, p<.001). So, again, doing something is better than doing nothing.

We also considered influences of instruction mode (online vs in-person) or instructor type (teaching assistant vs principal instructor). As for the instruction mode, we found generally higher response rates for students who attended classes online compared to their peers who received in-person instruction. The differences were of similar magnitude for survey sample, population, and a sample of all instructors who did not participant in this survey (average response rates for each group ranged from 49.3% to 49.6% for in-person, and from 60.8% to 66.5% for online courses); the proportion of online courses did not differ substantially between sample and population and was generally with 2% rather small. For that reason, we did not expect it to exert significant leverage on the models and did not control its influence statistically. We found lower response rates for teaching assistants than for principal instructors, but these differences are within the range of 2-4% (average response rates are between 45.9% and 49% for teaching assistants and between 48.9 and 50.6% for principal instructors). Teaching assistants accounted for 3.8% of the evaluated courses for which we gathered data in our survey, and for 15.2% in the group of courses where instructors did not participate. As teaching assistants represented only a small number of courses and differed with regard to response rates only marginally from principal instructors, we did not see a need to control for this influence either.

4. Conclusion

The results from this representative sample indicate that providing class time for completing course evaluations and reminding students to take the course evaluation surveys are especially robust strategies to improve response rates. Emphasizing confidentiality concerns and explaining the use of course evaluations were also associated with higher response rates but, possibly due to methodological limitations, they did not explain variability in response rates above and beyond the impact of the other strategies. In sum, utilizing *any* strategy to encourage student participation in course evaluation will improve response rates with setting aside course's time for completing evaluations and providing reminders being the most impactful strategies at Lehigh University.

5. References

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.

Sundstrom, E. D., Hardin, E. E., & Shaffer, M. J. (2016). Extra credit micro-incentives and response rates for online course evaluations: Two quasi-experiments. *Teaching of Psychology*, *43*(4), 276-284.