

Effects of Course Delivery Mode and Course Evaluation Mode on Student Ratings of Instruction

Debbie McGhee and Nana Lowell
March 2015

INTRODUCTION

This report summarizes student ratings of courses conducted at the University of Washington, Seattle (UWS) in Autumn 2014. These analyses were carried out at the request of the Academic and Student Affairs Committee of the Faculty Council of Academic Standards (FCAS). The committee was interested in whether face-to-face and online sections were perceived by students to be of equivalent instructional quality.

The report compares evaluation results by course delivery mode (face-to-face, hybrid, online) and mode of evaluation (paper, online) to address two main questions:

1. Are courses taught online rated differently than courses taught face-to-face?
2. Are courses evaluated online rated differently than courses evaluated using paper forms?

METHOD

Student ratings of instruction at UWS are carried out by means of the Instructional Assessment System (*IASystem*), developed and operated by the Office of Educational Assessment (OEA). OEA has provided course evaluation services to UW since the early 1970's using machine-readable optical mark evaluation forms. OEA recently has rebuilt the database application supporting course evaluation services in order to enable departments and instructors to choose to administer evaluations either online or via traditional paper forms. Using *IASystem*, departmental coordinators specify course section type (face-to-face, hybrid, online) and delivery mode (online, paper) at the time they request evaluations for individual courses; these indicators can be exported with student response for purposes of analysis.

The present study examined evaluation results for all course sections evaluated during Autumn 2014. We verified section type designations entered by departmental coordinators by comparing *IASystem* records against the printed Time Schedule.¹ Of the 249 sections initially identified as having been taught online, 132 were confirmed by Time Schedule records. The remaining 117 sections were recoded within *IASystem* as face-to-face. The final number of sections per course type was as follows: 4,319 were face-to-face sections, 72 were hybrid sections, and 132 were online sections (total $N = 4,523$). These records comprised the analytic dataset for the current study.

¹ Course type is entered in *IASystem* by departmental coordinators rather than being uploaded from Time Schedule records because the latter does not include a designation for hybrid courses.

RESULTS

Descriptive Statistics

All *IASystem* evaluation forms include a set of common items that can be used to compare ratings across courses, instructors, or academic terms. Four global items capture students' overall evaluation of the course; the aggregate of these item ratings is reported as the *global median*. Additional common items provide information to adjust global ratings for known biases and to compute an index of student challenge and engagement. The *adjusted global median* is a regression-based modification of the global median that takes into account class size, reason for taking the course, and expected grade. The *challenge and engagement index (CEI)* is a combined index of four items referencing the degree to which students were "challenged" by the course. Table 1 shows descriptive statistics for each of these measures over all course sections evaluated at UWS in Autumn 2014. Results are displayed by section type (face-to-face, hybrid, online) and evaluation mode (paper, online).²

Table 1. Course evaluation ratings by section type and evaluation mode (all courses)

Section type and evaluation mode	Global Median			Adjusted Global Median			CEI		
	Mean	SD	n	Mean	SD	n	Mean	SD	n
All face-to-face sections	4.1	.63	4,294	4.1	.59	4,188	4.8	.72	4,194
Face-to-face (paper)	4.2	.62	1,709	4.2	.58	1,668	4.8	.70	1,672
Face-to-face (online)	4.1	.64	2,585	4.1	.59	2,520	4.9	.73	2,522
All hybrid sections	4.1	.73	72	4.1	.69	66	4.8	.76	66
Hybrid (paper)	4.6	.35	9	4.5	.43	5	4.7	.71	5
Hybrid (online)	4.0	.75	63	4.0	.70	61	4.8	.77	61
All online sections	4.0	.57	131	4.0	.54	130	4.8	.73	130
Online (PCE)	3.9	.42	23	4.0	.39	23	4.5	.62	23
Online (self-sustaining)	4.0	.58	91	4.0	.56	90	4.9	.76	90
Online (IAS or Time Schedule)	4.0	.70	17	3.9	.61	17	4.9	.59	17
All paper evaluations	4.2	.62	1,718	4.2	.58	1,673	4.8	.70	1,677
All online evaluations	4.1	.63	2,779	4.1	.59	2,711	4.9	.73	2,713
All evaluations	4.1	.63	4,497	4.1	.59	4,384	4.8	.72	4,390

Comparisons by Section Type and Course Delivery Mode

We carried out a series of one-way analyses of variance to determine whether evaluation ratings (global median, adjusted global median, and CEI) were related to either section type (face-to-face, hybrid, online) or evaluation mode (paper, online).

² The number of ratings reported is slightly lower than the total number of completed evaluations due to missing data.

Section Type

Average global median and adjusted global median ratings were somewhat lower for online sections than for face-to-face or hybrid sections and, although the differences were small, analysis of variance confirmed that they were statistically significant. No difference was found for the CEI; students rated face-to-face, hybrid, and online sections as equally challenging. These results are summarized in Table 2.

Table 2. Analyses of variance comparing course evaluation ratings by section type

Section type	Global Median			Adjusted Global Median			CEI		
	Mean	SD	n	Mean	SD	n	Mean	SD	n
All face-to-face sections	4.1	.63	4,294	4.1	.59	4,188	4.8	.72	4,194
All hybrid sections	4.1	.73	72	4.1	.69	66	4.8	.76	66
All online sections	4.0	.57	131	4.0	.54	130	4.8	.73	130
	$F(2, 4494) = 3.78, p = .02$			$F(2, 4381) = 3.18, p = .04$			$F < 1$		

We considered that although the lower ratings observed for online courses might be due to a true difference in instructional quality, they also could originate from other factors such as a consistent student bias in favor of face-to-face or hybrid instruction, or differences in the type of courses that are offered online rather than only in face-to-face or a hybrid format. This suggested a third research question:

3. Are differences in ratings due to the *type of courses* which may be taught online (rather than solely due to differences in instructional quality)?

We began by determining whether there were differences in ratings of face-to-face sections of courses that also had online sections versus courses that did not have online sections. As shown in Table 3, analyses of variance indicated that face-to-face courses with an online alternative received lower global median ratings than did face-to-face courses without online equivalents. However, there was no difference in adjusted global median ratings. These results suggest: 1) that differences in ratings of online versus face-to-face sections may be due to the type of course for which online equivalents have been created, and 2) that this difference is corrected for by the adjusted global median.

Table 3. Analyses of variance comparing ratings of face-to-face courses with and without online equivalents

Online equivalent	Global Median			Adjusted Global Median		
	Mean	SD	n	Mean	SD	n
Had online equivalent	4.0	.67	269	4.1	.62	269
Did not have online equivalent	4.2	.63	4,025	4.1	.58	3,919
	$F(1, 4290) = 21.0, p < .001$			$F(1, 4184) = 1.88, p = .17$		

To further explore whether differences in ratings were due to the type of course evaluated, we next restricted our analysis to only those courses which had both face-to-face and online

sections. As shown in Table 4, there were no significant relationships between section type and either global median or adjusted global median ratings.

Table 4. Analyses of variance comparing ratings by section type among courses with online equivalents

Online equivalent	Global Median			Adjusted Global Median		
	Mean	SD	n	Mean	SD	n
Face-to-face sections	4.0	.67	269	4.1	.62	269
Online sections	4.0	.54	113	4.0	.52	113
	$F < 1$			$F(1, 380) = 2.42, p = .12$		

Evaluation mode

Sections evaluated online were given somewhat lower global median and adjusted global median ratings than were sections evaluated using paper forms. Additionally, sections evaluated online were rated as somewhat more challenging than were sections rated on paper. These results are summarized in Table 5.

Table 5. Analyses of variance comparing course evaluation ratings by evaluation mode

Section type	Global Median			Adjusted Global Median			CEI		
	Mean	SD	n	Mean	SD	n	Mean	SD	n
All paper evaluations	4.2	.62	1,718	4.2	.58	1,673	4.8	.70	1,677
All online evaluations	4.1	.63	2,779	4.1	.59	2,711	4.9	.73	2,713
	$F(1, 4495) = 5.76, p = .02$			$F(1, 4382) = 18.9, p < .001$			$F(1, 4388) = 8.44, p = .004$		

CONCLUSION

Over all sections, global ratings, but not the CEI, were somewhat lower among sections delivered online than among those with face-to-face instruction. However, these differences dissipated when the analyses were restricted to only those courses that offered an online alternative. These results suggest that when examining whether there are differences in perceived course quality by delivery format, it is important to parcel out extraneous course characteristics which may affect ratings; that is, to compare only "like to like." Courses with online offerings tend to have larger enrollments and are more likely to be lower-level courses, and these factors have been shown to be negatively related to global evaluation. Furthermore, because students give more critical ratings via the internet compared to in class, online sections are susceptible to this additional complication.

The slightly lower ratings obtained via the internet rather than via paper also may reflect some type of bias rather than true differences in instructional quality. Possibilities here include greater perceived (rater) anonymity for online evaluations, and being outside the "halo" of the classroom. These could be tested by comparing online and paper evaluations when both are administered under the same conditions, for example, in-class ratings of face-to-face courses.