University of Washington
Faculty Council on Teaching and Learning
April 2nd, 2015
10:30am - noon
Gerberding 142

Meeting Synopsis:

1) Call to order
2) Review of the minutes from March 5th, 2015
3) Analysis of course evaluation data from last autumn (Nana Lowell / Debbie McGhee) (Exhibit 1)
4) Guide for Tenure and Promotion evaluations - input from Council (B. Kalikoff)
5) Course Evaluation Catalogue: resolution to stay online (Exhibit 2)
6) Reports from FCTL Subcommittees (Exhibit 3)
7) Planning for our annual report (Wilkes)
8) Additional items (members please propose to JW)
9) Adjourn

---

1) Call to order

Wilkes called the meeting to order at 10:30 a.m.

2) Review of the minutes from March 5th, 2015

The minutes from March 5th, 2015 were approved as written.

3) Analysis of course evaluation data from last autumn (Nana Lowell / Debbie McGhee) (Exhibit 1)

Nana Lowell (Director of the Office of Educational Assessment) and Debbie McGhee (Research Scientist, Office of Educational Assessment) reported they had analyzed data from autumn 2014 with the intention of answering two questions pertaining to course evaluations conducted online, versus conducted in the classroom on paper. The research was requested by the Faculty Council on Academic Standards (FCAS). The two questions to be answered were:

1. Did ratings collected on paper differ from those collected via the internet?
2. Did ratings of face-to-face sections differ from ratings of internet-based sections?

In autumn quarter, there were 2800 courses evaluated online, and 1700 on paper. McGhee explained the analyses also took into account whether a student took the course in-person, online, or in a hybrid course.

Conclusions drawn from the data

The guests explained that a number of conclusions were drawn from the data. The conclusions were (listed in Exhibit 1):

OFFICE OF EDUCATIONAL ASSESSMENT
Improving Learning Through Assessment

OEA Report 15-02

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 a 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

	Global Median			Adjusted Global Median		
Online equivalent	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

	Global Median			Adjusted Global Median			CEI		
Section type	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.

- Global ratings were somewhat lower among sections delivered via the internet compared to those taught face-to-face.
- However, the differences dissipated when the analyses were restricted to only those courses that offered an online alternative.
- In addition, ratings given via the internet tended to be more critical than those given on paper, regardless of instruction mode.
- Taken together, the results suggest that group differences in ratings may reflect response biases (e.g., rater anonymity, type of course), rather than true differences in instructional quality.
  ** Full report available at: http://www.washington.edu/oea/pdfs/reports/OEAReport1502.pdf*

Council feedback

Turner expressed interest in further specific analysis of online course evaluations in the future. He noted he would like to see one evaluation type used predominantly at the UW, rather than both (online and in-person) used interchangeably.

Wilkes noted he is surprised that the differences in course ratings (shown in graphs, Exhibit 1) seem so minor, given the anonymity of the internet and the casualness of internet ratings. Turner explained he would prefer sample-size numbers to be shown along with the data for better understanding. He noted if the council was considering making weighty decisions over online learning, there should be more data pertaining to student satisfaction and instructor effectiveness. Lowell explained that she and McGhee did not have enough data to answer several additional questions, and explained the autumn quarter 2014 was the first time they had a very large data pool for analysis. Turner explained he finds that the most capable members of his department are among those conducting courses online.

Lowell explained 60% of course evaluations were offered online in autumn 2014, with the remaining paper evaluations accounting for 40%. The council expressed interest in knowing the ratios of paper versus online evaluations in past years – in looking at trends in the university moving all evaluations online.

Wilkes explained the obstacles in evaluating faculty in any measurable way. He noted he would like to see a thorough report looking at deeper correlations between evaluation data and various factors affecting class delivery. He added as a council, it would be beneficial to have ideas for specific studies of the data, for researching topics that the council may find to be of importance. He asked that the council submit any ideas to him for possible deeper analysis. He then noted he would like to form a working group, to meet before the next meeting, and discuss possible ideas for a topic they might have interest in delving deeper into.

Lowell explained that this particular report was especially and purposely brief. McGhee is working on additional analyses in looking at faculty and student demographics. Wilkes suggested that graduate students may be able to help in this effort.

Wilkes explained the council would come back to this topic if areas of interest could be identified.

4) Guide for Tenure and Promotion evaluations - input from Council (B. Kalikoff)

Beth Kalikoff (Director, Center for Teaching & Learning) explained the Center for Teaching and Learning (CTL) is in the process of drafting a best practices guide for assessing teaching to be used on campus for tenure and promotion committees. Kalikoff explained that the proposed timeline for the guide would be that it would be drafted in winter quarter 2015, and then go to a review process including various faculty bodies - with an aim to be revised, finalized, and ready for use by tenure and promotion committees in the fall of 2015. The project was authorized and requested by (then) Provost Ana Mari Cauce. Kalikoff noted the CTL has no authority or desire to force groups on campus to use or adhere to the best practices outlined in the guide. She explained she hoped the guide may be useful for tenure and promotion committees and others for assessing teaching, and spare groups from conducting the research themselves. She explained the guide hopes to provide options for discretion in departments for assessing teaching.

The guide is six pages long, and includes 3-4 additional pages of sources and additional material of use. Kalikoff explained the guide is divided into three parts. The parts are:

I. Self-assessment – A self-critique of an instructor's effectiveness. It was noted several literary sources boast this method to be the most effective.
II. Peer review & observation
    a. Observation of teaching: witnessing an instructor's teaching in the classroom, or participating in their online course
    b. Assessment of assignments: evaluating all the supporting materials that comprise the work of the course
III. Students' data – quantitative evaluation data (e.g. course evaluations)

Kalikoff noted that in their research, they found that each and all departments use student evaluations, and nearly all use some form of peer review. It is known many faculty members do not understand self-review, what it is, and what it looks like. Kalikoff thanked the Office of Educational Assessment for their enormous contribution to the project.

She noted that the most challenging practice, the one that causes the most uneasiness with faculty, is the assessment method of peer review. She explained this phenomenon occurs because norms often vary greatly between teaching styles, and questions linger over universally valued ways and styles in which faculty may teach. She explained that when observing a classroom, there is marked uncertainty in what the observer is looking for.

Kalikoff compared teaching assessment to a "three legged stool." She noted one of the reasons they parse the data is because student ratings have such a high impact. She noted one of the most implicit arguments of the guide is that student assessment are very important, but, it still is undecided if they are more important than peer review and/or self-assessment. She explained realigning and redistributing the stakes must be considered.

Kalikoff clarified that her and her group's base assumption, is that the faculty at the UW represent a very exceptional group of instructor's. The data and guide is not intended to be used to penalize faculty, but to identify and promote best practices of teaching, which can be located through the research.

Kalikoff noted she would like feedback from the council on any additional groups the CTL should visit for feedback, as well as advice on the order of which they should be consulting these groups (which groups first, which before others, etc.) for the greatest efficiency. She also noted that they would like to

understand if it is best to make changes to the guide as they collect feedback, or to come with a final draft of the best practices to all groups.

Wilkes began by noting self-assessment is not well defined. Kalikoff agreed and noted the document only offers a rubric for self-assessment, and provides links for other information and rubrics from other institutions. She explained the guide offers examples from peer institutions and from within the UW on how to self-evaluate. She added it is a contested area of conversation, and there is no singular response for how to evaluate in this manner. She explained they are also looking specifically at what students are best at evaluating.

Spyridakis asked if the document states what its purpose should be, or what the purpose of the document is. Kalikoff explained the purpose, though she was not sure if the document's purpose was explicitly stated within the document itself. Spyridakis advised caution in moving forward with the document, in avoiding misunderstandings.

Kalikoff noted the document will be distributed the week after this council meeting, and will be discussed at the next meeting. A member noted the document reads very well, and they were impressed with its overall readability. Wilkes noted he would like to see numbered revisions, for showing when sections were changed and by whose suggestion. Moreover, Wilkes noted he would suggest the document be presented as the final version, and not be formally revised by each body reviewing it. Turner noted there are standing promotion and tenure committees in the foster school of business which the document may be useful for.

Wilkes noted he hopes this will become an online document in the future which may be accessed seamlessly. Kalikoff noted any revisions will be made to an online copy.

The council thanked Kalikoff for updating the council on the effort.

5) CEC resolution to stay online (Exhibit 2)

Wilkes noted there are problems with funding and data scraping with the Course Evaluation Catalogue (CEC), which will result in it being taken down. The council had taken action on the CEC being replaced by the MyPlan alternative, discussed in the last three council meetings. Wilkes noted he would not like these problems to be the reason the CEC is taken offline, as it is important for students. He has drafted a Class C resolution for the intention of beginning discussion over keeping this resource online for the good of students (Exhibit 2).

Wilkes noted the goal is to press resources to be used for keeping the CEC online.

Lewis noted he will find more details on the problems associated with the CEC and report back to the council. Lowell explained that they have the data from the CEC archived and stored; it is just a matter of posting the data online to make the evaluations available for other parties (students, faculty, etc.). She noted the CEC was originally developed by UW-IT, and is currently 15 years old. She noted the data is not deleted annually, or at any other point. Wilkes explained his main proposal is to make the CEC available again, but to keep it in a harder "protective shell" – protected from scraping, and from other misuses.

Wilkes explained discussion on this item will continue.

Proposed resolution: Keeping CEC accessible online
JW, 1st draft 4/1/15

Context:

The Course Evaluation Catalog (CEC) has long been an important resource for students (and faculty). The proposed 3-item boxes on MyPlan can only report results from the latest offering of a given course by a given instructor. This does not begin to provide the depth and breadth of information on faculty evaluations CEC provides. For example, students in large-enrollment introductory courses, where instructors often change from term to term, will certainly want to explore the evaluations of the instructor they will have.
The CEC will be taken offline because its database has been vulnerable to 'scraping' by outside organizations, especially commercial 'rating' websites that typically present raw evaluation data without context or explication.

It does not seem possible that there is simply no solution to the problem of restricting access to the CEC to UW faculty and currently-active students.

Of course UW-IT has better understanding of the difficulties, but perhaps this could be done by replacing text with images, or adding CAPTCHA or similar challenge-response protections to safeguard against access by automatons or unauthorized users.

One concern is that any authorized user can publish the data. However, the council could (for example) ask FCAS to propose legislation that would make it an academic offense for persons with authorized access to publicly post evaluation data, or knowingly pass it on to unauthorized persons.

Resolution:

The council urges UW-IT to investigate ways of keeping CEC data accessible to legitimate users. It is acceptable for such access to be moderately inconvenient e.g. having several security steps), since that is far preferable to having access eliminated except by special arrangement with OEA.

6) Reports from FCTL Subcommittees

a) DL subcommittee report update (Wilkes, R. Corbett) (Exhibit 3)

Wilkes noted he would like the council to review the document for the purpose of jointly endorsing the FCAS guideline, and for possibly advising any additional changes.

Wilkes noted a DL course is defined as a course that is almost completely conducted online, where a student would almost never have to be present on campus to participate. He explained courses which have substantial online offerings, but hold "regular in-person meetings," are not defined as DL within the document anymore; this is a substantive change.

Wilkes noted one issue that came up concerns the residency requirement, and the fact that DL courses do not count towards overall residency at the university. He explained the subcommittee has recommended keeping the three-year course review for DL courses, as they find it necessary for ensuring quality, among other reasons.

Robert Corbett (President, Professional Staff Organization) noted Tina Miller in the Registrar's Office had clarified that the state of Washington wants to know how many classes are taught online in its universities. He explained that the understanding was if a department wants to declare a course DL, than they can choose to do so. Wilkes noted the state reporting requirement was formally of concern, though the state only wants it reported the amount of courses which are taught online, and the data does not see a high amount of analysis by any means.

Wilkes noted he would like the council to read the Subcommittee on Distance Learning's report, and advise any changes, or make any comments in the next council meeting.

7) Planning for our annual report (Wilkes)

Wilkes noted at the next meeting he would like to present an annual report of the FCTL's activities for the 2014-2015 academic year. He noted if outside work has been done, it should be reported in or before the next council meeting.

8) Additional items (members please propose to JW)

Kalikoff noted she would like to update the group on a joint center for teaching and learning in the next meeting, briefly.

9) Adjourn

Wilkes adjourned the meeting at 11:47 a.m.

Minutes by Joey Burgess, jmbg@uw.edu, council support analyst

Present: **Faculty:** Jan Spyridakis, Jennifer Taggart, Dan Turner, Jeffrey Wilkes (chair)

Faculty Council on Academic Standards Policy on DL (distance learning) designation for UW Seattle courses

A course or section of a course in which students can participate fully without being physically present on campus must be designated as a "DL" course or section. This includes courses in which some, but not necessarily all, offerings of the course are online, such as courses that are offered with different technologies in different quarters (i.e., one quarter on campus and another quarter online) and courses where, in a quarter, one section of the course is on-campus and another section is online. A DL course may have occasional meetings on campus for the purposes of organization, evaluation, or group presentations.

However, courses that make extensive use of online teaching tools but also have regular recurring meetings (i.e., weekly), such as labs, discussion sections, studios or other on campus offerings, do not typically require the DL designation, even if these meetings are for shorter duration than those traditionally encountered in on-campus courses.

The DL designation is obtained via the course creation/change process. Courses designated DL do not count for residency and must be re-approved after 3 years, in accordance with Student Regulations Chapter 115.1.I.

Approved by the Faculty Council on Academic Standards March 20, 2015

Ex-officio reps: Terry Ann Jankowski, Hailey Badger, Eldridge Alcantara, Robert Corbett
Guests: Rovy Brannon, Nana Lowell, Christine Sugatan, Tom Lewis

Absent:

Faculty: David Masuda, Ellen McGough, Bruce Nelson, Jaime Olavarria, Brenda Zierler
Ex-officio reps: N/A
President's designee: Ed Taylor

Exhibits

Exhibit 1 – Course Evaluations: Online vs. Paper evaluations
Exhibit 2 – CEC Resolution
Exhibit 3 – DL Subcommittee Report