Report on results of the Online Pilot of the Carleton Teaching Evaluations (March/April 2012)

The pilot of an online teaching evaluation (TE) system was conducted this winter term (2012) under the auspices of an agreement between CUASA and the university – the agreement text is available at: http://oirp.carleton.ca/teforms/Memorandum-of-Agreement-Teaching-Evaluation-Pilot.pdf

Pilot participation:

The online pilot was conducted during the last two weeks of class in the winter term 2012. In total, 143 CUASA (and related) members participated in the pilot although one instructor did not have a valid matching course for comparison. Students were invited to participate in the online pilot by an email invitation to their 'Connect' email – students were sent up to three email reminders to participate in the pilot if they didn't complete their online evaluations. The bulk of students had one evaluation in the online pilot while a small number had multiple evaluations to a maximum of four. There were 7,398 students involved in the online pilot. Instructors volunteering for the TE pilot were asked by email and paper distribution to encourage their students to participate in the online teaching evaluations for their course by clicking on the link in their connect email (or email system where they had forwarded their connect email). An additional reminder was sent asking instructors to further encourage participation and this email included an indication of the TE participation rate for their course.

The link in students' email forwarded them to a 'Teaching Evaluation Portal' which indicated the course or courses that they could complete online. A click on the course link took them to a similar form to the paper evaluation – the two forms currently used were available as appropriate: one for FASS, FPA and Business and one for Science and Engineering (attached below).

The breakdown of participating instructors/sections is as follows:

Online Teaching Evaluation Pilot - Number of Participating Instructors by Faculty and Level

	Level					
	1	2	3	4	5	All Levels
Arts & Social Sciences	13	15	14	9	5	56
Public Affairs	4	5	6	6	13	34
Science	7	2	6	6	2	23
Engineering & Design	3	6	6	2	5	22
Business	1	2	2	2	1	8
All	28	30	34	25	26	143

Source: Teaching Evaluation Pilot March/April 2012

There is a good mix of participants across faculties and levels though, given the total number of participants, individual cell sizes (faculty/level combinations) are generally quite small making it difficult to draw conclusions about differences at this detailed level.

Participation rates:

As expected response rates fell somewhat between the paper and online evaluation – this has been the experience across institutions that have switched to an online system. In general the decline in participation rates has been 8% on average though in two cases the decline was more precipitous. The median response rate decline was 6.25. The following histogram shows the change in response rates:

Mean = -8.12
Std. Dev. = 19.802
N = 142

Change in Response Rate by Course (Online Rate - Paper Rate)
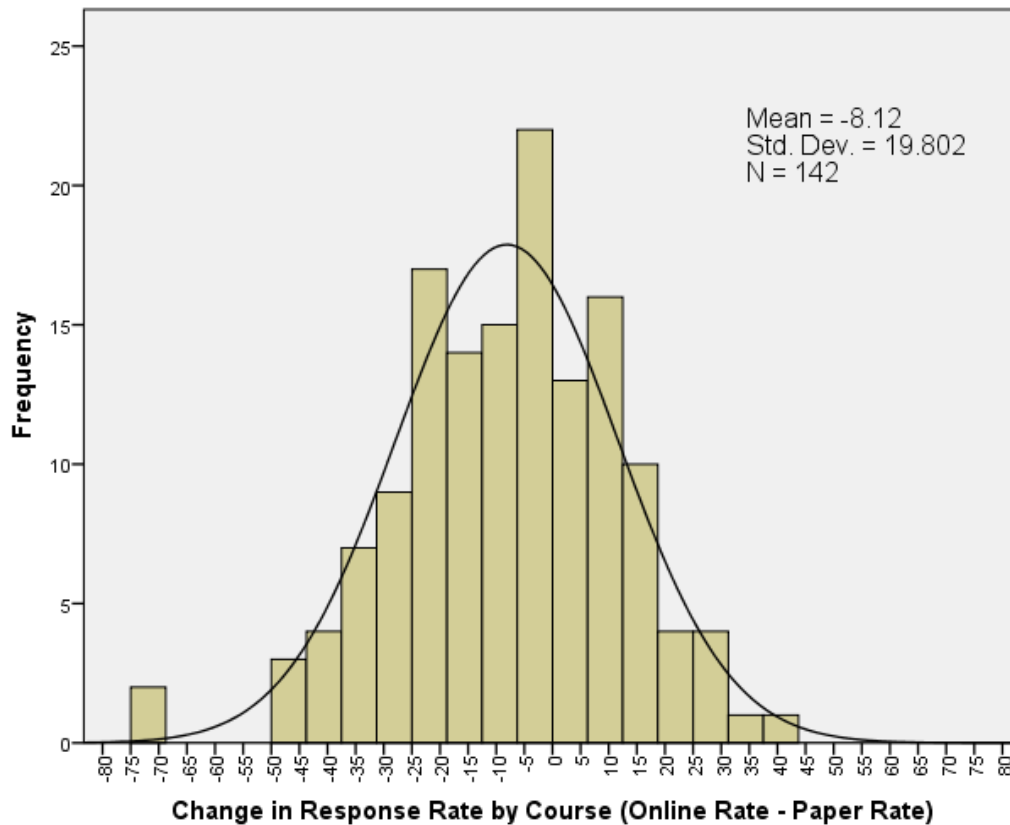
Since we track access to the online teaching evaluation system (though we do not track individual responses for confidentiality reasons), we know that following the closing of the evaluation system on the last day of classes there were students still trying to access the system. In fact a further 5% tried to fill out their evaluations in the four days following the closing of the system to responses. We interpret this as a failure to adequately communicate the appropriate dates for filling out the form and the lag in students' reference to their 'Connect' email (we think that some number of students only consult their Carleton email on an irregular basis and some may not have checked until after evaluations were closed for the term). In a full scale system we would be able to activate an alternative method of connecting to the evaluation system through login using the Carleton MyPortal credentials – this method would facilitate greater advertisement of the open evaluation system as notices regarding the online evaluation system could contain a link that would take students automatically to the evaluation portal.

Response rates do not have a noticeable pattern by faculty and level and it is important to take note of the small 'N' for some cells in the table (see participating Instructors above).

Student response rates by faculty and level, for the pilot, are included below:

Online Teaching Evaluation Pilot - Comparison of Average Response Rates (%)
by Faculty and Level

| | Level | | | | | | | | | | All Levels | |
| | 1 | | 2 | | 3 | | 4 | | 5 | | | |
	Online	Paper	Online	Paper	Online	Paper	Online	Paper	Online	Paper	Online	Paper
Arts & Social Sciences	51	66	49	57	55	63	62	78	86	82	56	66
Public Affairs	37	48	41	30	49	71	61	73	71	87	57	69
Science	39	43	62	66	57	58	67	60	53	71	54	56
Engineering & Design	53	62	51	55	62	57	52	70	65	71	58	62
Business	55	67	55	60	66	70	60	96	71	82	61	75
All	46	57	49	53	56	63	62	74	71	81	57	65

Source: Teaching Evaluation Pilot March/April 2012
Notes: 1. The online course evaluation is matched to the latest paper based evaluation for the same instructor (e.g. from Winter 2011).

In general response rates are highly variable depending mainly on the level of the course – last year's winter evaluation section response rates were:

Average Section Response Rates for Teaching Evaluations by Faculty and Level
(Winter Term, 2010/11)

| | Lev. 0 | | Lev. 1 | | Lev. 2 | | Lev. 3 | | Lev. 4 | | Grad. | | All | |
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Arts & Social Sciences			135	61.0%	175	59.8%	151	62.9%	110	76.2%	64	85.9%	635	69.2%
Public Affairs			71	60.2%	59	49.6%	94	65.9%	110	69.7%	116	84.4%	450	66.0%
Science	3	26.8%	46	40.0%	36	50.0%	43	58.2%	41	72.2%	20	77.9%	189	54.2%
Engineering & Design			23	47.1%	39	51.7%	43	57.9%	60	65.6%	40	73.9%	205	59.2%
Business			11	56.5%	29	62.2%	26	68.8%	28	81.9%	25	86.2%	119	71.1%
All	**3**	**26.8%**	**286**	**53.0%**	**338**	**54.7%**	**357**	**62.8%**	**349**	**73.1%**	**265**	**81.6%**	**1598**	**63.6%**

Source: Aggregate Teaching Evaluations
Notes: 1. N -> number of sections
2. % -> average section response rate.

<u>Average Scores:</u>

Average scores did not vary in a statistically significant way between the online and previous paper evaluations. A repeated measures analysis of variance with or without faculty and level 'between subject' factors did not reveal any statistically significant differences. Average scores were used for comparison since this is the best method of comparing across the two separate teaching evaluation forms – average scores are the average response to questions 1 through 12 for FASS, FPA and Business and 1 through 13 for Science and Engineering. The analysis of variance results follow (a 'Sig' greater than .05 indicates a lack of statistical significance in the means at the 95% level of confidence):

Tests of Within-Subjects Contrasts with Level and Faculty

Source	online_paper	Type III Sum of Squares	df	Mean Square	F	Sig.
online_paper	Linear	.010	1	.010	.148	.701
online_paper * level	Linear	.067	4	.017	.260	.903
online_paper * faculty	Linear	.101	4	.025	.389	.816
online_paper * level * faculty	Linear	1.179	16	.074	1.136	.331
Error(online_paper)	Linear	7.588	117	.065		

Tests of Between-Subjects Effects (Faculty and Level)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	3209.700	1	3209.700	21006.572	.000
level	.756	4	.189	1.237	.299
faculty	.677	4	.169	1.107	.357
level * faculty	2.422	16	.151	.991	.472
Error	17.877	117	.153		

Tests of Within-Subjects Contrasts (Using Online/Paper only in Model)

Source	online_paper	Type III Sum of Squares	df	Mean Square	F	Sig.
online_paper	Linear	.161	1	.161	2.541	.113
Error(online_paper)	Linear	8.917	141	.063		

A graph of the difference in the average score (Paper average – Online average) shows that in general the differences in scores tend to be small though there are a few outliers where the differences are larger – some to the high side (i.e. online average is greater than the paper average) and some to the low side. There are two evaluations with unexpectedly large differences between the online and paper version – the exact cause is, of course, unknown, but both of these courses have low response rates and, in absolute terms, few responses (in one case only 2 responses). It is, of course, important to maintain a representative sample from a given section to ensure a reliable result – this appears to have occurred in the vast majority of courses but there appears to have been some special circumstances or lack of communication that impacted a few courses. The histogram of average score differences follows:
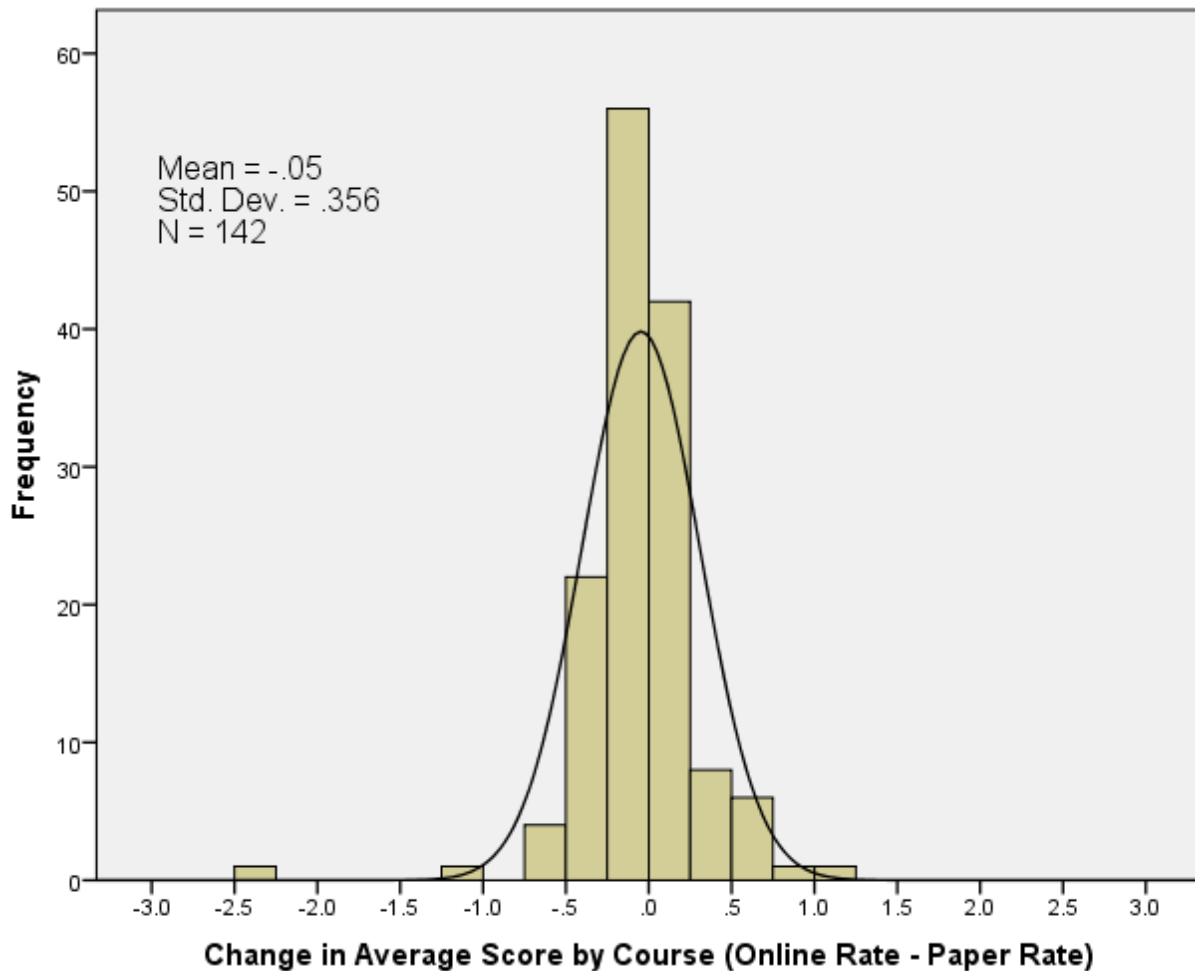
The overall difference in average scores was -.0476 and the median was -.0455. The trimmed mean (ignoring the biggest and lowest 5% of cases) was -.0433.

A scatterplot of differences between Online and Paper show that the bulk of evaluations cluster about the 'x=y' line indicating that the online and paper version of the surveys have similar scores – those below the diagonal will have a higher online score (than paper) and those above will have a lower score. The two low outliers are evident in this graph as well as a few high outliers.

Comparing Paper and Online Average Evaluation Score

N = 2 (of 9)

N = 14 of 47

<- Online Average = Paper Average

Online Response Rate
○ <= 42
○ 43 - 48
○ 49 - 55
○ 56 - 63
○ 64 - 73
○ 74+
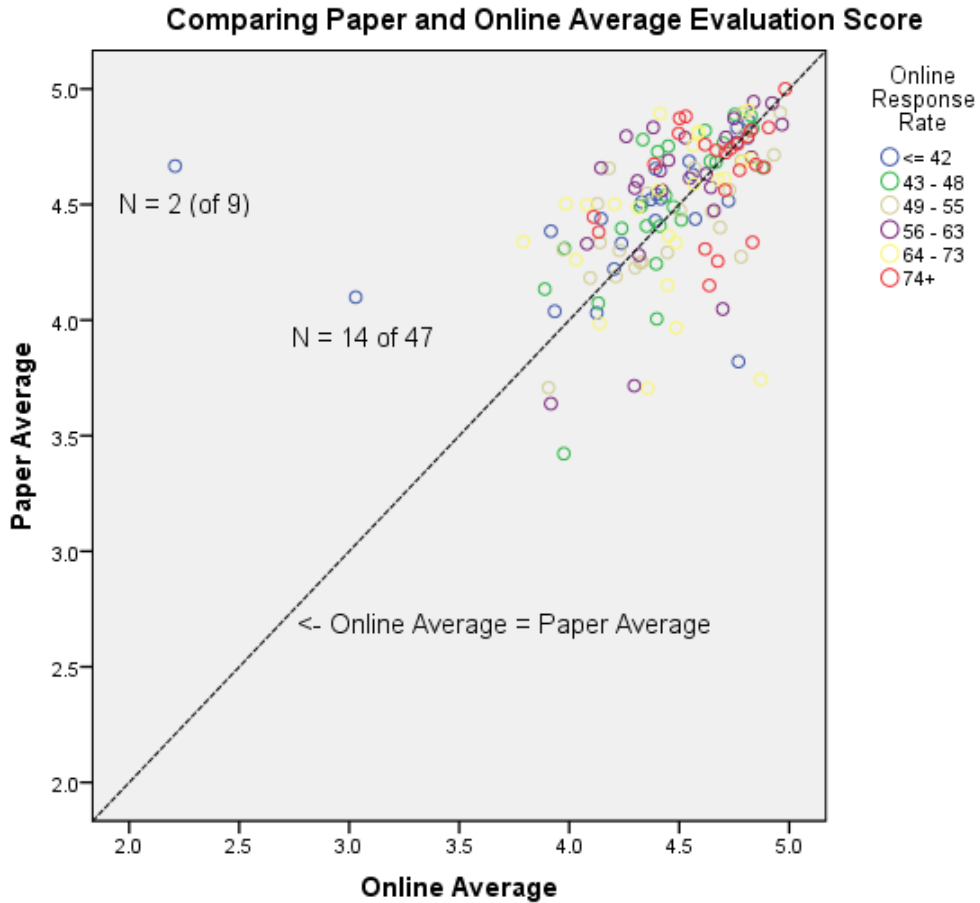
Average scores by level and faculty follow for the overall average and for Question B (for FASS, FPA and Business) – overall differences are generally in the second decimal place.

Online Teaching Evaluation Pilot - Comparison of Average Results by Faculty and Level

	Level										All Levels	
	1		2		3		4		5			
	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.
Arts & Social Sciences	4.62	4.57	4.39	4.40	4.25	4.51	4.58	4.62	4.66	4.68	4.46	4.53
Public Affairs	4.32	4.35	4.36	4.48	4.10	4.52	4.73	4.75	4.63	4.60	4.48	4.57
Science	4.21	4.53	4.77	4.72	4.58	4.47	4.43	4.43	4.66	4.44	4.45	4.50
Engineering & Design	4.32	4.26	4.25	4.34	4.49	4.52	4.45	4.49	4.53	4.70	4.41	4.48
Business	4.30	4.23	4.37	4.49	4.47	3.93	4.26	4.39	4.45	4.37	4.37	4.28
All	4.43	4.48	4.38	4.43	4.34	4.47	4.54	4.58	4.61	4.61	4.45	4.51

Online Teaching Evaluation Pilot - Comparison of Overall Question Results (Question B)
by Faculty and Level (for FASS, FPA, and Business)

	Level										All Levels	
	1		2		3		4		5			
	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.	Online Avg.	Paper Avg.
Arts & Social Sciences	4.64	4.56	4.31	4.37	4.22	4.49	4.52	4.64	4.71	4.77	4.43	4.52
Public Affairs	4.30	4.33	4.29	4.47	3.90	4.45	4.74	4.82	4.67	4.64	4.45	4.58
Business	4.30	4.26	4.29	4.37	4.48	3.96	4.27	4.39	4.39	4.32	4.35	4.25
All	4.54	4.49	4.30	4.39	4.15	4.43	4.57	4.67	4.67	4.66	4.43	4.52

Source: Teaching Evaluation Pilot March/April 2012

Notes: 1. The online course evaluation is matched to the latest paper based evaluation for the same instructor (e.g. from Winter 2011).

Some interesting facts:

Since browsers 'announce' themselves when they connect to a website we can track the origin of accesses to the teaching evaluation system. In total there were 4841 accesses (there may be multiple accesses for a completed evaluation) to the system and of these (numbers may not add to 100% due to rounding):

Operating Systems/platform:

- Windows: 66%
- Macintosh: 29%
- Smartphone: 7%
- Linux: 1%

Smartphone type:

- Apple devices (Iphone, Ipod, Ipad): 72%
- Android: 16%
- Blackberry: 12%

Browser:

- Webkit based (Chrome/Safari etc.): 60%
- Firefox: 23%
- Internet Explorer: 16%
- Other: 1%

Windows based browsers (62% of windows systems are 64 bit):

- Chrome: 47%
- Firefox: 27%
- Internet Explorer: 24%
- Other: 1%

<u>Next Steps:</u>

In the opinion of OIRP the pilot was a success and we recommend that we move to the full online trial for this summer. That is, all summer evaluations (two evaluation periods) be conducted online – this will largely impact contract instructors as relatively few full-time faculty teach in the summer. Last summer 27% of evaluated sections were taught by CUASA members and most of the rest by contract instructors.

This summer would provide us with the opportunity to run a full scale online evaluation and, assuming we do not discover any critical problems, then we would recommend the adoption of a full-scale online system for the fall of 2012.

Some observations:

- Average scores did not vary on aggregate in a statistically significant way from the paper to the online version.
- Response rates did decline but, in general, not in a magnitude that would put the measurement of the representative scores at risk.
- Maintaining a 'reasonable' response rate in each section is important (to the paper and online versions). In the paper version the response rate is fixed as to the attendance in the class on the date the evaluations are undertaken. In the online version the response rate is not given and depends mostly, we think, on encouragement given to fill out the online form by instructors in class and the general level of advertisement of the system.
- We think that a full scale system, as compared to the pilot, will offer better opportunity for advertising and making the online teaching evaluation system accessible to students. We could use an alternate method of access which allows students to login to the system without using the link provided in their email invitation. The email invitation provides a simple access methodology but requires students to actually READ their university email in a timely manner. The login method allows us to better advertise the evaluation period throughout the university web-site and through the learning management system by imbedding an active link to the TE system.
- The response rate for the pilot is excellent for an online survey without 'incentives' (an incentive being some sort of prize associated with filling out the survey). In addition we had substantial numbers of students trying to access the system after the closing date for the evaluation – we think this is a failure to communicate effectively the dates for the survey (that is differentiating between communication and effective communication). This is clearly an area that can be improved in the next round.
- Smartphone access constitutes 7% of accesses to the system – while not currently large; this is an area to which we need to pay attention as it will likely grow in time. Modifying the system to be Smartphone (i.e. small screen) aware would be a nice feature but is technically challenging due to the necessity to ensure absolute security of access and not allow students to fill in more than one evaluation. To accommodate Smartphone use we think that the extra request for confirmation when a student clicks the submit button is confusing and is the cause of some incomplete responses (currently, at the request of CUASA, we have programmed a further query

into the system when the user clicks submit – they are prompted to confirm their request to submit). We would recommend removing this extra check as we can expect that students are fully familiar with web forms and will not click the submit button in error.

- Concomitant with this analysis of the online pilot we realized that the high speed scanner which is used to evaluate the paper version of the teaching evaluations has had a serious malfunction. Interim plans have been put in place to complete the winter evaluations albeit with some days delay. It reinforces the fact that the current system is dependent on older hardware/software technology that is past due for a complete reinvestment. We would prefer to move to an online system so resources can be effectively employed to improve and modernize the system not maintain technology that is a decade and a half old.
- The online system functioned well during the pilot and no serious technical errors were noted.
- There were a number of positive communications with faculty as to the benefits of the online system. In particular, a number of faculty commented on the class time saved near the end of the year as a result of using the online system. One comment was received regarding the number of communications we had with instructors (one on paper and two emails).

Carleton
UNIVERSITY
Canada's Capital University

Teaching Evaluation Questionnaire

Teaching Evaluation: BUSI1402H - Alejandro Ramirez

Course: BUSI1402H - Intr Business Info & Comm Tech
Instructor's Name: Alejandro Ramirez

Information for Students:
1. This evaluation is confidential and anonymous and is only released to the Instructor after grades have gone to the Dean.
2. This questionnaire is used by the university to evaluate the performance of teaching staff. It may be used in decisions affecting the Instructor's promotion, tenure, or salary increase.
3. Comments may be entered in the space provided at the end of this questionnaire. They will be forwarded to the Instructor after the final grades have been processed (only the Instructor will read your comments).

[Start]

Carleton
UNIVERSITY
Canada's Capital University

Teaching Evaluation Questionnaire

Teaching Evaluation: BUSI1402H - Alejandro Ramirez

Course: BUSI1402H - Intr Business Info & Comm Tech
Instructor's Name: Alejandro Ramirez

1. **Please answer all of the following questions with a rating between 1 (Poor) and 5 (Excellent), by selecting the appropriate bubble.**

How do you assess your instructor's performance:	1 - Poor	2	3	4	5 - Excellent	Not Applicable
in making clear the objectives of the course?	○	○	○	○	○	○
in organizing the course so as to meet the objectives?	○	○	○	○	○	○
in imparting the course materials in his/her role as /lecturer /seminar leader/tutorial leader/workshop leader/language instructor?	○	○	○	○	○	○
in answering questions and/or solving problems related to the course material?	○	○	○	○	○	○
in assigning readings, essays, seminar topics, etc., which are relevant to the course?	○	○	○	○	○	○
in assigning a workload related to the course objectives?	○	○	○	○	○	○
in marking and commenting on assignments and tests fairly?	○	○	○	○	○	○
in returning tests and assignments promptly?	○	○	○	○	○	○
in being available for out-of-class consultation?	○	○	○	○	○	○
in speaking audibly and clearly?	○	○	○	○	○	○
in beginning and ending classes promptly?	○	○	○	○	○	○
in meeting classes regularly as scheduled, and in missing or cancelling classes only for adequate reasons?	○	○	○	○	○	○
How do you evaluate the instructor?	○	○	○	○	○	○

2. **Comments may be entered in the space provided below. They will be forwarded to the instructor after the final grades have been processed. Only the Instructor will read your comments.**

[Submit]

Sample Online Evaluation Form 2

Carleton
UNIVERSITY
Canada's Capital University

Teaching Evaluation Questionnaire

Teaching Evaluation: AERO3002A - Jeremy Laliberte

Course: AERO3002A - Aerospace Design and Practice
Instructor's Name: Jeremy Laliberte

Information for Students:
1. This evaluation is confidential and anonymous and is only released to the Instructor after grades have gone to the Dean.
2. This questionnaire is used by the university to evaluate the performance of teaching staff. It may be used in decisions affecting the Instructor's promotion, tenure, or salary increase.
3. Comments may be entered in the space provided at the end of this questionnaire. They will be forwarded to the Instructor after the final grades have been processed (only the Instructor will read your comments).

[Start]

1. **Please answer all of the following questions with a rating between 1 (Poor) and 5 (Excellent), by selecting the appropriate bubble.**

   How do you assess your instructor's performance:

	1 - Poor	2	3	4	5 - Excellent	Not Applicable
in making clear the objectives of the course?	○	○	○	○	○	○
in organizing the course so as to meet the objectives?	○	○	○	○	○	○
in preparing for class?	○	○	○	○	○	○
in imparting the course material to the students?	○	○	○	○	○	○
in answering questions and/or solving problems related to the course material?	○	○	○	○	○	○
in making assignments which are relevant to the course?	○	○	○	○	○	○
in assigning a workload related to the course objectives?	○	○	○	○	○	○
in being available for out-of-class consultation?	○	○	○	○	○	○
in speaking audibly and clearly?	○	○	○	○	○	○
in beginning and ending classes promptly?	○	○	○	○	○	○
in meeting classes regularly as scheduled and in missing or cancelling classes only for adequate reasons?	○	○	○	○	○	○
in marking and commenting on assignments and tests fairly?	○	○	○	○	○	○
in returning tests and assignments promptly?	○	○	○	○	○	○

2. **Comments may be entered in the space provided below. They will be forwarded to the instructor after the final grades have been processed. Only the Instructor will read your comments.**

[Submit]