A Comparative Analysis of the Student Course Perception Survey (SCP) & Faculty Course Evaluation Data for the 2018 Fall Term

August 18, 2021

Abstract

The aim of this comparative analysis is to investigate the following two areas of interest: (1) How much volatility should instructors expect in the transition between the new Student Course Perception (SCP) survey and Course Evaluation (CE) instruments currently used at the end of the term (i.e. the old metrics); and (2) Does the move to the new instrument either improve or exacerbate gender differences in scores? To answer these questions, pilot test responses from the Fall 2018 SCP survey were compared with end-of-term Fall 2018 CE survey results for each faculty. The results from this comparative analysis reveal the following: Though some movement in scores across instruments is to be expected, generally instructors will do roughly as well on the new instrument, as compared to their colleagues, as they do on the old instrument. Instructors were very likely to remain in the same quartile across both instruments and when movement occurred it was very likely to involve movement up or down only one quartile. Moreover, cases where we saw movement in scores beyond one quartile in either direction were almost always in classes with a low number of responses. Another key finding from this analysis is that the transition from the existing CE metrics to the SCP instrument does not seem to change how average scores from female instructors compare to their male colleagues, although a clear trend was difficult to decipher. In some Faculties, the proportion of female instructors with scores above the median, was slightly higher for the SCP than for the CE, while in others it was slightly lower, but in all Faculties differences are not statistically significant.

Contents

1 Introduction

As part of the implementation of the pilot test for the new Student Course Perception (SCP) survey carried out in the Fall 2018, the Course Evaluation (CE) Project Team held consultations with designated "point people" (often Undergrad Associate Deans or Teaching Fellows) representing each Faculty. The point people from some Faculties urged the view that many instructors would value some comparison between the results of the new SCP instrument and the existing CE survey instruments in use in their courses. There were two general questions that were judged to be of interest: (1) How much volatility should instructors expect in the transition between instruments (e.g., instructors might worry, 'if I am performing very well with the old CE tool, how great is the risk that my scores will plummet with the new tool?'). And (2) Does the move to the new SCP instrument either improve or exacerbate gender differences in scores? To answer these questions, pilot test responses from the Fall 2018 pilot SCP survey are compared with end-of-term Fall 2018 CE survey results for each faculty that elected to participate in this analysis. Faculties that elected to participate signed an agreement to share course evaluation data from the Fall 2018 term.

The purpose of this report is to examine the questions outlined above while highlighting some methodological decisions made in light of the challenges involved when comparing the old CE metric to the new SCP metric. All Faculties were offered the opportunity to participate in this comparative analysis. Four Faculties opted to do so: Engineering, Mathematics, Health (formerly AHS) and Science. It is worth noting, the Health and Arts instruments are the same, but appropriate caution should be exercised when using the evidence described below to judge what we might have observed if we had done an analysis of the Arts data.

2 The Data Set

2.1 Student Course Perception Survey Questions

Currently, at the end-of each term, each Faculty utilizes their own set of questions to evaluate the performance of instructors. Most questions are measured on a 1-5 point scale, where scores are ranked from poor to excellent. There are also some text-based questions which are not included in the analysis of this report. A complete list of survey questions for each Faculty, is included in the Appendix.

In contrast, for the pilot test of the new SCP instrument, the same set of questions were asked across each Faculty. These items are measured on a 5-point Likert scale and include the following:

1. The instructor identified the intended learning outcomes for this course.
2. The intended learning outcomes were assessed through my graded work.
3. The course activities prepared me for the graded work.
4. Graded work was returned in a reasonable amount of time.
5. The instructor helped me to understand the course concepts.
6. The instructor created a supportive environment that helped me learn.
7. The instructor stimulated my interest in this course.
8. Overall, I learned a great deal from this instructor.
9. Overall, the quality of my learning experience in this course was excellent

In order to make comparisons across the old CE surveys and new SCP instrument, it was necessary to omit some items assessed in the Fall 2018 pilot test. For example, students' own perception of the course workload and self-reported course attendance rate are used as independent variables as opposed to dependent variables like the above metrics in the SCP pilot test regression analysis. For further details, please see the Course Evaluation Project Pilot Test — Data Analysis Report.

2.2 Descriptive analysis

Figure 1 provides an overview of the sample including the number of students in each Faculty who completed a regular end-of-term course evaluation, and the number of courses included from each Faculty (in brackets). For example, in the Engineering Faculty 17447 students submitted regular end-of-term course evaluations from 330 courses. Similarly, the number of students and the number of distinct courses for the 2018 SCP survey are shown in Figure 2.

3 Course-level Comparative Analysis

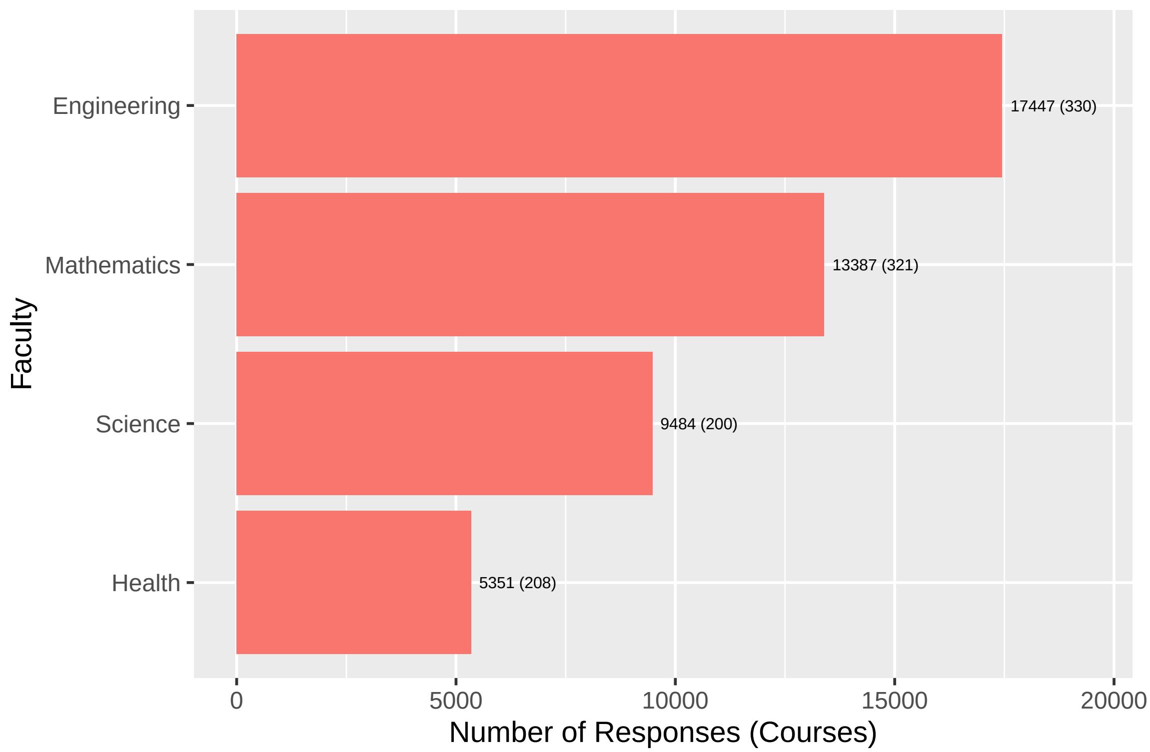

Figure 1: Numbers of students and courses (in brackets) by Faculty for the Fall 2018 Course Evaluation (CE) survey

3.1 Comparison Principles

While the same set of questions was tested in every Faculty for the pilot test, each of the four Faculties that opted to participate in the comparative analysis currently uses a different survey instrument (see the appendix for the questions used in each Faculty). We judged that in these circumstances attempting campus wide comparisons, or comparisons of particular Faculty results for the Course Evaluation instrument to campus results to the SCP instrument made little sense. We therefore, split the pilot test data set into Faculty specific data sets. Comparisons in this report are made within each Faculty.

Given that course evaluation survey responses remain anonymous for students, for this analysis we did not have student-level identifiers in the data set. Therefore, we could not link the student submissions from the SCP pilot test survey to the regular end-of-the-term CE survey. Comparisons in this analysis are made at the course-level, across course-instructor pairs (CIPs). In this sense, course A taught by instructor X and course B taught by the same instructor X constitute two distinct CIPs. In other words, we would have a score for Instructor X course A, and a second score for Instructor X course B (same instructor, different courses). The precise definition of the aggregate score we use for the SCP and CE instruments on each CIP is detailed in Section 3.3.1.

3.2 Responses for course-instructor pairs

Course-instructor pairs (CIPs) were used to calculate the number of responses for each instructor. This includes the sum of total responses (excluding NA responses) in each question column for each CIP. See Figure 3 for the valid number of CIPs by Faculty. One of the questions investigated was potential volatility in scores, which we anticipated would be more likely in courses with a low number of responses.

A *valid response* to any evaluation question for a course is a student response with a meaningful score choice (e.g. 1 - 5 points as required). Items left blank were coded as NA (invalid) and were deleted from the analysis.

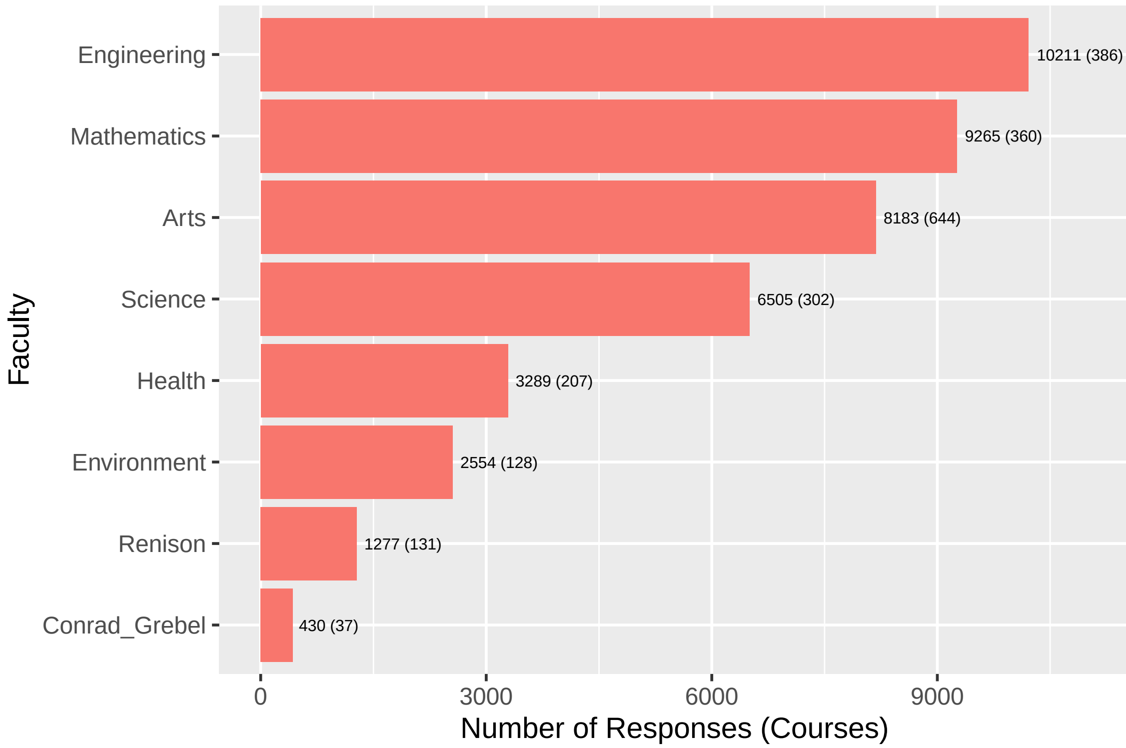

Figure 2: Number of students and courses (in brackets) by Faculty for the Student Course Perception (SCP) Survey.

Preliminary data cleaning revealed the sample included many CIPs with 10 or fewer valid responses for some course evaluation items. Further analysis revealed that these low response counts explained plot outliers (i.e., volatility in scores).

To examine the effect of valid student responses (i.e., sum scores excluding NA responses) on the average course evaluation score for each CIP, we consider the *number of responses* per CIP. This is defined as the *minimum* number of valid responses across all the survey items for each CIP. Suppose for a specific CIP, the numbers of valid responses for question 1, ... question n on the Course Evaluation survey (old metric) are a_1, ... a_n; similarly the numbers of valid responses for question 1, ... question m in the SCP pilot test survey (new metric) are b_1, ... b_m, then the minimum number of valid responses for this CIP is:

$$N_{\text{min}} = \min\{a_1, \ldots, a_n, b_1, \ldots, b_m\}$$

This computation allows us to classify CIPs into three groups based on the number of valid responses they received: 0-5, 6-20, 20+. Figure 4 shows the decomposition of the number of responses by Faculty.

By comparing Figure 1, Figure 3 and Figure 4, we can see that this calculation results in a reduced sample size.

3.3 Comparison of Instructor Rankings

Before conducting the analysis, one key issue to address is how to compare the evaluation results between the 2018 Course Evaluation (CE) metric and the pilot SCP survey, given that they do not use the same items to assess students perceptions of instructor performance. In this section, we explain our procedures for computing a comparable rank score for both evaluation metrics. Results are compared at the Faculty level.

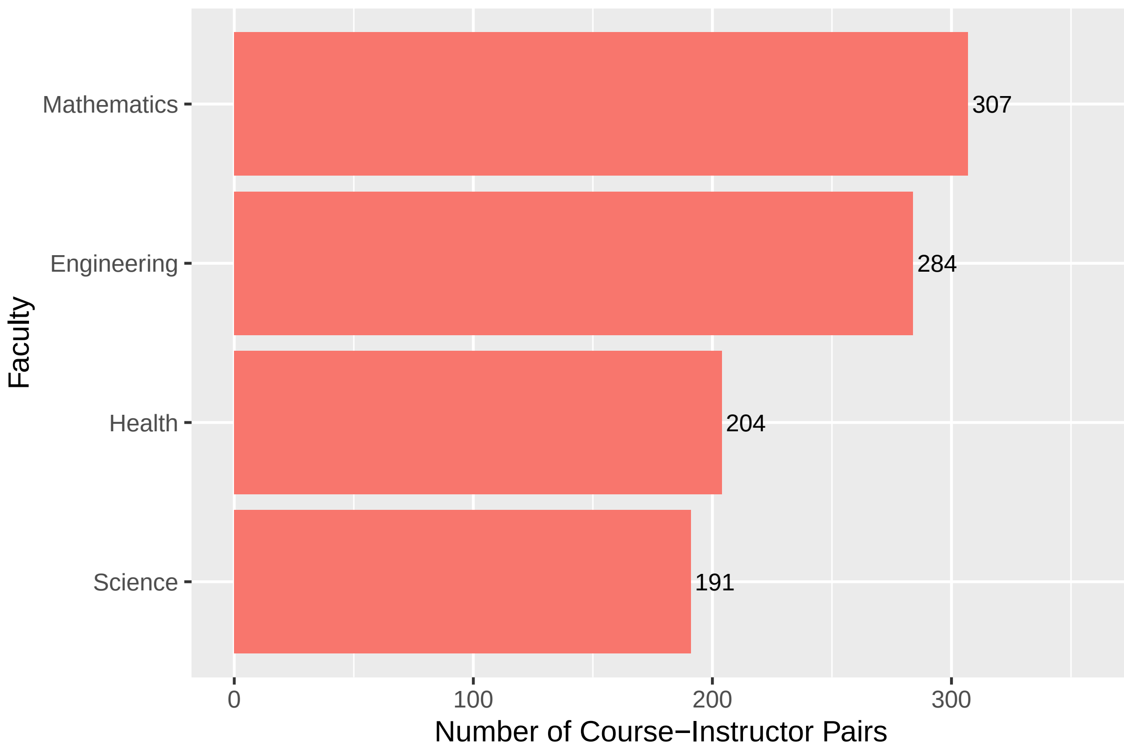

Figure 3: The number of course-instructor pairs (CIPs) by Faculty. Invalid CIPs with NA responses or unmatched pairs between Course Evaluation and SCP data sets are removed.

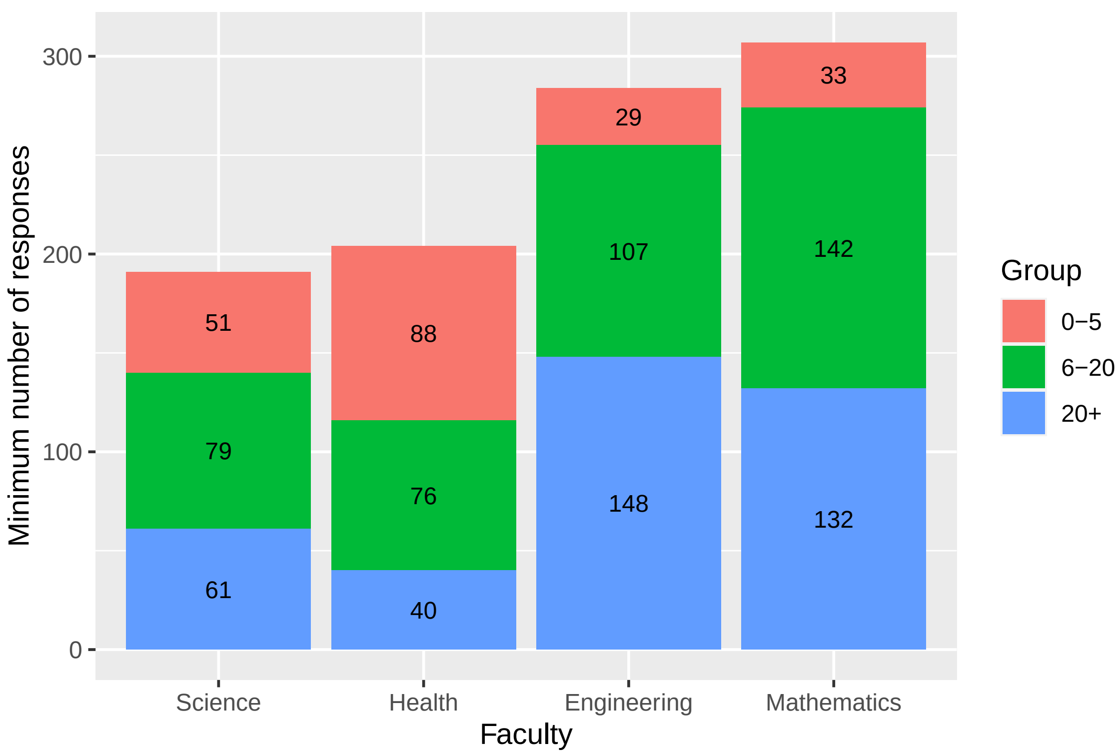

Figure 4: The number of responses by Faculty.

3.3.1 How to compute the aggregate score

To obtain an aggregate evaluation score for each instrument in each CIP, we simply compute the average of all scores (typically on a 5-point Likert scale) across all evaluation questions and students in the given CIP.

To allow for comparisons to be made across instruments, it was necessary to exclude some survey items from the old CE surveys in each Faculty. Specifically, we excluded items that:

- were optional for students (e.g. describe the availability of the instructor outside of class);
- had high instances of missing data (aka N/A responses);
- were not obviously related to an instructor's performance (e.g. workload or measures of study hours, these factors are explanatory variables in the SCP pilot test regression model);
- were open-ended (usually text-based).

After excluding these items, average scores for the existing instruments (old CE metric) were calulated using the items below for each Faculty (please see the Appendix for a list of the items included):

- For Health, questions 0 ~ 8 are included;
- For Engineering, questions 0 ~ 3, 5 ~ 10 and 13 ~ 16 are included (question 5 and 10 are coded so that the middle-scale option is the 'highest'(comparable to a score of 5 on the 5-point Likert scale used for the SCP pilot study), question 4 was optional and thus omitted as it had several missing cases);

- For Mathematics, questions 0-7, 9, 12,13 were included, and questions 1,12,and 13 were recoded so that the middle item is the 'highest'(comparable to a 5 on a 5-point Likert scale);
- For Science, questions 3 ~ 9 are included.

3.3.2 Calculating comparisons across two different survey instruments

It is important to emphasize that methodologically, comparing average numerical scores from surveys that ask different questions makes little sense. More prosaically, the new SCP survey measures items on a 5-point Likert scale, while several Faculties currently subscribe to different scales. With this in mind, in coding the items for this analysis we made every effort to address these shortcomings and elected to use *rank score* (explained in detail below) to allow for consistent comparisons across different evaluation metrics. Visually, we adopt rank scatter plots to display our findings (Please see Figure 5). In the figures that follow, the Student Course Perception pilot test survey is titled "SCP rank" and the end of term Fall 2018 course evaluations (CE) are labelled "CE rank."

Rank scores are calculated by ranking the aggregate CIP scores from lowest to highest within a given Faculty and then converting this to a quantile between 0 and 100. In other words, if a given CIP has a rank score of 75, then it's score is higher than 75% of CIPs in the same Faculty.

Figure 5 depicts scatterplots of SCP vs CE rank scores within each Faculty. Each point is colored according to number of minimum valid responses per question as defined by N_{min} above. The red dots correspond to CIPs with 0-5 responses, green dots represent CIPs with 6-20 responses and finally, blue dots represent CIPs with 20+ responses.

These plots reveal that most points (especially the blue and green points) are spread out along (fall close to) the 45 degree diagonal line which suggests instructors who perform well on the old CE metric are likely to perform well on the new SCP metric. There are some points which are a considerable distance from the diagonal line which means there is a substantial discrepancy between the old and new scores for these particular CIPs. It is worth noting, a large porportion of these dots (for instance, in the case of Health almost all of them) are red which means we are dealing with extremely small numbers of valid student responses (0-5). Caution is warrented in attributing too much weight to these findings given the extremely small sample sizes we are dealing with. These findings are consistent with our pilot-test results which caution against giving significant weight to SCP scores when there is a low number of student responses.

To further illustrate this point, the broken lines (—) on each plot indicate a band inside of which SCP and CE rank scores differ by no more than 25%.

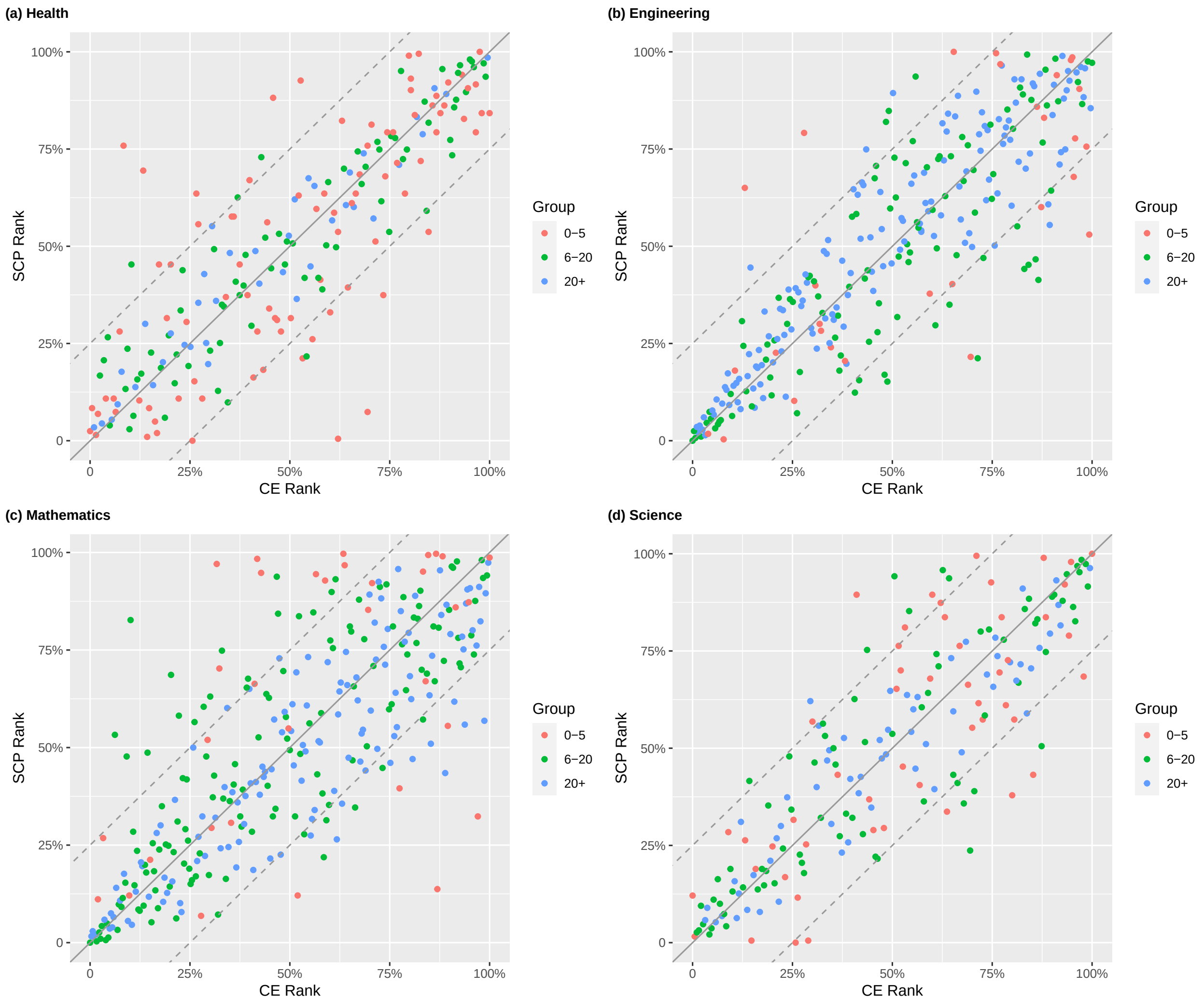

Figure 5: New SCP rank vs old CE rank by the minimum response numbers for all four Faculties. Higher rank is better. Each dot represents a course-instructor pair (CIP). The CIPs are categorized into three groups based on their number of responses: 0-5, 6-20, 20+

The data reveal a clear pattern for nearly all the points that fall outside of the 25% tolerance region defined above (i.e., the outliers), these points are associated with 20 (or fewer) responses at the minimum level (as signified by the red and green dots). As noted above, the most extreme outliers (red dots) have minimum response numbers <5. This pattern tells us that when a CIP has a low number of responses, the comparison between the old and new metric is unreliable (in other words, scores have the potential to change dramatically). Of course, with a low number of responses, greater volatility is to be expected, which is why SCP results with very low numbers of responses need to be interpreted with caution.

3.4 Assessing Changes in Rank Quartiles

In this section, the extent to which a CIP might see a large jump or decline in their average score when moving from the old CE metric to the new SCP metric is examined.

The tables below compare the quantile range for scores from the old CE metric and the new SCP metric. For example, results reveal that in Health there were 204 CIPs in total and thus, 51 CIPs in each quantile. Of those 51 CIPs that fell into the 0-25% quantile range on the old metric, when compared to the new metric, 37 remained in the 0-25% quantile, 12 moved to the 25-50% quantile, 1 moved to the 50-75% quantile and 1 jumped to the 75%-100% quantile. For the 51 CIPs who fell in the 75-100% quantile on the old metric, 42 remained in this quantile on the new metric, while 9 moved to the 50-75% quantile for the new metric and no one fell below the 50-75% quantile on the new metric.

Table 1: Contingency table for ranking quartile changes for all Faculties. Each row represents a group ranked by the old CE metric. Each column is a group ranked by the new SCP metric. Higher rank means better evaluation.

	SCP metric				
CE metric	0-25%	25%-50%	50%-75%	75%-100%	total
Health					
0-25%	37	12	1	1	51
25%-50%	10	27	13	1	51
50%-75%	4	12	29	6	51
75%-100%	0	0	9	42	51
total	51	51	52	50	204
Engineering					
0-25%	56	14	1	0	71
25%-50%	13	40	15	3	71
50%-75%	2	13	39	17	71
75%-100%	0	4	17	50	71
total	71	71	72	70	284
Mathematics					
0-25%	58	15	3	1	77
25%-50%	16	34	22	5	77
50%-75%	2	23	27	24	76
75%-100%	1	5	24	47	77
total	77	77	76	77	307
Science					
0-25%	38	10	0	0	48
25%-50%	9	25	12	2	48
50%-75%	1	11	20	15	47
75%-100%	0	2	15	31	48
total	48	48	47	48	191

For the other Faculties (Engineering, Mathematics and Science), we can observe the following general pattern:

- Instructors were very likely to remain in the same quartile across both instruments;

- When movement occurred it was very likely to involve movement up or down only one quartile.

3.5 Assessment of Gender Bias

The other question of particular concern to Faculty representatives is whether the move to a new instrument will either exacerbate or mitigate any instructor gender differences in the scores. We address this question using the framework of statistical hypothesis testing. That is, we begin with the null hypothesis:

H_0: Instructor gender differences are the same in the new instrument as the old one.

Note: This is not the same thing as assuming that there are no gender differences in either instrument. Rather, it is stating that whatever difference exists must be the same in both instruments. We then examine the statistical significance of the differences in gender gaps between the two instruments using the metrics described below.

3.5.1 McNemar's test

Given the limitations imposed by the data available, we approached this question as follows:

Under H_0, we expect that the number of CIPs with female instructors above the median CIP score with the CE instrument is the same as the corresponding number with the SCP instrument.

We can then examine the porportion of female CIPs above the median for each of the instruments to see the extent to which female CIPs change when moving from the old CE metric to the new SCP instrument. To examine this we used McNemar's test to explore the potential for statistically significant gender differences when the old CE metric is compared to the new SCP survey. For this test, "above median" represents scores that fell above the 50% quantile, or equivalently, having a rank score above 50. The results for this test for each Faculty are presented in the contingency tables below.

Table 2 shows the proportion of female CIPs above the median for each instrument, and about how much movement in scores occurred across the median. For example, for female CIPs in Health, using the old instrument, 47 were below the median and 57 were above. For the new instrument, 46 were below and 58 were above. Of the 47 female CIPs who scored below the median on the old metric, 39 remained below the 50% quartile, while 8 CIPs scored "above the median," on the new metric. Conversely, of the 57 CIPs who scored "above the median" on the old metric, 50 remained "above the median" and 7 moved to "below the median" when evaluated on the new metric.

The p-values for all Faculties in Table 2 are well above 5%, indicating that there is little evidence that gender differences change from the old instrument to the new one. In other words, we could say that the gender difference in each of the two instruments is roughly the same.

3.5.2 Wilcoxon signed-rank test

McNemar's test metric – the difference in the number of female-instructor CIPs above the median between the old instrument and the new one – is relatively easy to understand. However, McNemar's test is not especially powerful, in that it cannot detect small differences in the gender gap between the two instruments without large sample sizes. In contrast, the Wilcoxon signed-rank test is a much more powerful test against H_0. The drawback is that the test metric is more difficult to interpret (and indeed we do not attempt to do so here).

The Wilcoxon test also assumes that the difference between rank scores on the new and old instrument has a symmetric distribution. These differences are plotted in Figure 6, which reveals that this assumption is by and large correct. Figure 6 also includes p-values for the two-sided and one-side Wilcoxon tests. The one-sided test checks whether female instructors fare worse relative to their male colleagues under the new instrument than the old. The two-sided test only looks for differences in gender gap between the two metrics, without distinguishing for which metric it is better or worse. In all cases, the p-values are well above 5%, indicating little evidence of a difference in gender gap between the two instruments.

Table 2: McNemar's test for all Faculties. Each row represents a group under the old CE metric. Each column is a group under the new SCP metric. We should not reject the null hypothesis at a significance level of 0.05, i.e. there is no strong evidence that the change in instruments will change how female instructors score relative to male instructors.

	SCP metric			
CE metric	Below Median	Above Median	total	p-value
Health				
Below Median	39	8	47	
Above Median	7	50	57	0.796
total	46	58	104	
Engineering				
Below Median	30	6	36	
Above Median	5	28	33	0.763
total	35	34	69	
Mathematics				
Below Median	30	4	34	
Above Median	8	28	36	0.248
total	38	32	70	
Science				
Below Median	28	1	29	
Above Median	5	23	28	0.102
total	33	24	57	

4 Concluding Remarks

In this report, our analyses reveal the following:

1. Generally instructors will do roughly as well on the new SCP evaluation metric, as compared to their colleagues, as they do on the old evaluation metric. Instructors were very likely to remain in the same quartile across both instruments and when movement occurred it was very likely to involve movement up or down only one quartile. A key finding, is that in cases where we saw significant movement across quantiles, these could be explained by CIPS with extremely low valid responses.

2. The transition from the existing Course Evaluation metric to the new SCP metric does not seem to change how average scores from female instructors compare to their male colleagues. There was no clear pattern evident with respect to the change in scores for CIPs. In some Faculties female CIP scores appeared to improve with the transition to the new instrument (e.g. an increase of one quartile), while for other Faculties we observed the opposite for some female CIP pairs (e.g., an average score falling in one quartile lower). Regardless of the shift in scores, in all cases, the differences are not statistically significant. Therefore, more data are needed to conclude that whatever differences we see are not merely due to random chance.

5 Appendix: List of questions for each faculty

We include a list of the course evaluation questions asked in each Faculty survey. In Applied Health, questions 0 – 8 are measured on a 1 – 5 point scale, where scores are ranked from poor to excellent. Text-based questions (10 – 12) are removed from this study. Question 9 measures workload and therefore is also excluded from our analysis. See the Table 3 below for the full list of questions.

For Engineering (undergrad), in Table 4, we recode questions 0 – 3, 6 – 9 and 12 – 16 to align with a typical 1 – 5 Likert scale. For questions 5, 10, 11, we recode each item as follows: the middle score (C) is assigned the highest score (5), so that it is equivalent to a score of 5 on the SCP likert scale. The extreme scores on

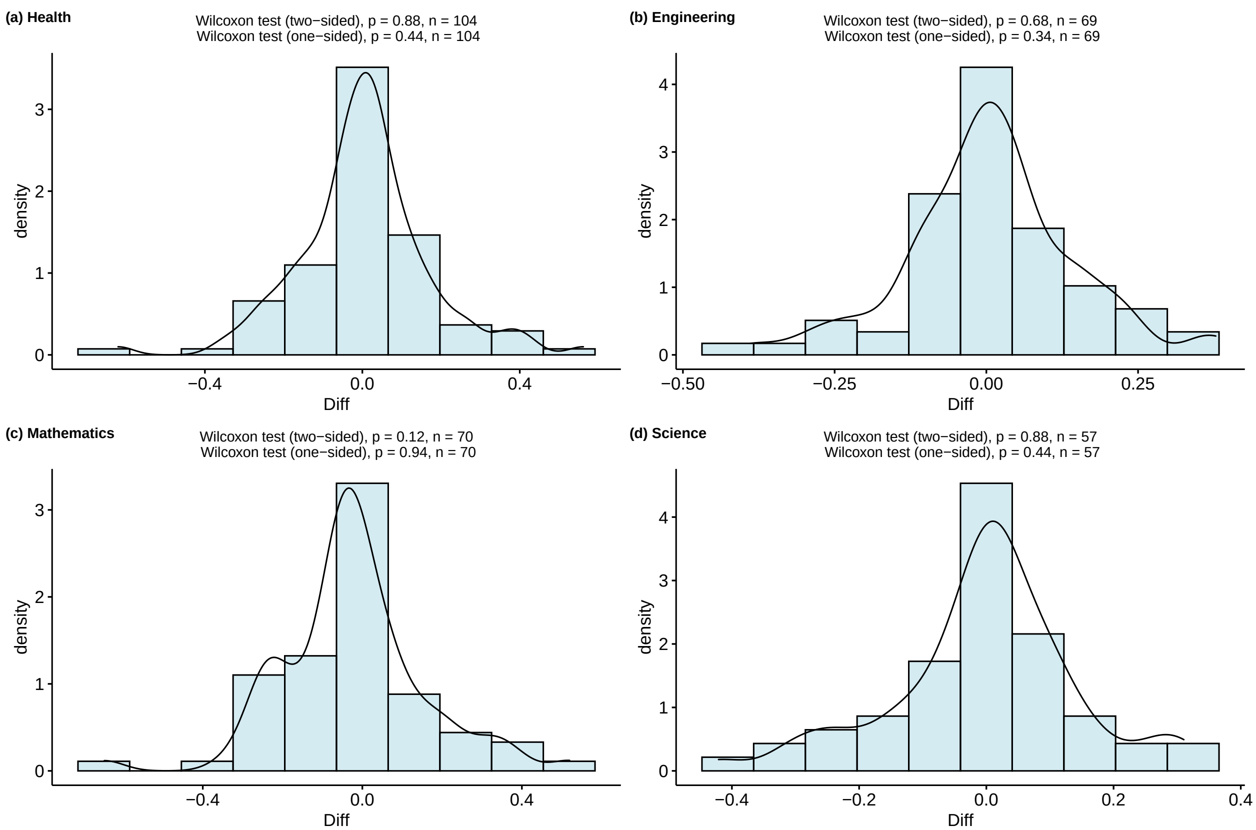

Figure 6: Wilcoxon signed-rank test for the female course-instructor pairs (CIPs) with histograms for the difference between the rank scores of the female CIPs under the new SCP metric and the old CE metric. The difference is new SCP score minus old CE score. For the one-sided test, the null hypothesis is that the female CIPs perform better under the new SCP metric than under the old CE metric; for the two-sided test, the null hypothesis is that there is no significant difference between the two metrics. At a significance level of 0.05, we reject the null hypothesis for p-value less than 0.05.

Table 3: List of questions for Applied Health

Question Index	Question Text
0	Rate the course with respect to the presentation of course/lab material:
1	Rate the course with respect to ability to maintain student interest:
2	Rate the course with respect to course/lab organization and planning:
3	Rate the course with respect to instructor's attitude toward and interest in students:
4	Rate the course with respect to instructor's availability outside of class:
5	Rate the course with respect to overall evaluation of the instructor:
6	Rate the course with respect to objectivity and fairness in discussions and grading:
7	Rate the course with respect to value of readings, assigned work, lab reports
8	Rate the course with respect to overall evaluation of the course/lab:
9	Rate the course with respect to workload demands upon the student:
10	In my opinion the strengths of this course/lab are:
11	In my opinion the weaknesses of this course/lab are:
12	My view of this course/lab overall is:

the original scale (A, E) are re-coded as 1 while (B, & D) are recoded as 3. Question 4 is excluded from the analysis, given it is optional and has a high number of missing cases).

Table 4: List of questions for Engineering (Undergrad)

Question Index	Question Text
0	Rate your professor's presentation of lectures in terms of organization and clarity.
1	How would you rate the professor's response to questions?
2	Rate your professor's oral presentation in terms of audibility, articulation, and your ability to understand his/her English.
3	Rate your professor's visual presentation. Consider organization, legibility and effective use of materials (blackboard, electronic media, etc.).
4	Describe the availability and approachability of your professor outside of class compared to other courses you've taken at university. (Do not answer if you did not seek help.).
5	At what level were the professor's explanations usually directed, in relation to your general level of understanding? *Note that (C) is the best rating
6	Were students encouraged by the professor to think and reason logically and objectively on their own, about the subject matter?
7	What was your impression of your professor's attitude towards teaching the course?
8	Describe your professor-class relationship?
9	What is your overall appraisal of the quality of teaching in this course?
10	Rate the difficulty of the concepts covered by this course. *Note that (C) is the best rating
11	Rate the workload required to complete this course. *Note that (C) is the best rating
12	How useful was the textbook (or reading materials) for understanding of the concepts?
13	To what extent did the assignments contribute to your better understanding of the concepts.
14	How well did the tests reflect the course material?
15	How valuable do you feel tutorials are (or would be) in the instruction of this course?
16	What is your overall appraisal of this course?
17	How many classes did you attend?
18	Your professor would like to know if there is something you believe he or she has done especially well in his or her teaching of this course.
19	Your professor would also like to know what specific things you believe might be done to improve his or her teaching of this course.

For Mathematics, course evaluation questions are included in Table 5. We include questions 0 − 13 but omit items 5 and 8. For questions 1, 12 and 13, the middle score is assigned the highest value (5). Question 6 is recoded as follows: Very Insteresting (coded as 5), Insteresting (coded as 3), Not Interesting (coded as 1) and No opinion (coded as NA).

Table 5: List of questions for Mathematics

Question Index	Question Text
0	Evaluate the organization and coherence of the lectures:
1	At what levels were the instructor's explanations aimed?:
2	Evaluate the instructor's treatment of students' questions:
3	Evaluate the effectiveness of the instructor's visual presentation (blackboard, overheads, etc.):
4	Evaluate the effectiveness of the instructor's oral presentation:
5	Was the instructor available for help outside of class?:
6	Did you find the course interesting?
7	Evaluate the overall effectiveness of the instructor as a teacher:
8	What proportion of lectures did you attend in this course?
9	Was the assigned work (assignments, projects, etc.) helpful in learning the course content?
10	Were the printed notes (if any) helpful in learning the course content?
11	Was the required textbook (if any) helpful in learning the course content?
12	Did the course introduce an appropriate amount of new material?
13	Was the amount of assigned work required for the course appropriate?
14	On average, how many hours per week did you spend on this course outside of lectures?
15	Please mention anything that you feel the instructor has done well in this course.
16	Please make constructive comments about anything in the instructor's technique or style that could, in your opinion, be improved
17	What were the strong points of the course?
18	What were the weak points of the course?
19	Was the class atmosphere affected either positively or negatively by attitudes of the instructors or students, e.g., with respect to gender, race, appearance? Please explain.
20	Any other comments, e.g., class size, suitability of room, noise level, etc.

Finally for Science, (Table 6), we include questions $2-9$ and apply the same coding framework as previously described. Questions 0 and 1 are omitted from this analysis given that they ask about level of difficulty and workload.

Table 6: List of questions for Science

Question Index	Question Text
0	Rate the RELATIVE DIFFICULTY of this course compared to the others you are now taking:
1	Rate the LOAD OF ASSIGNED WORK in this course compared to others you are now taking:
2	Rate the HELPFULNESS OF THE TEXT (or other assigned reading material) relative to those in other courses you are now taking:
3	The MID-TERM TESTS were a fair test of the material covered:
4	Relative to the other courses I have taken this term, I have found the EDUCATIONAL AND/OR PROFESSIONAL VALUE of this course:
5	Rate your professor's ORAL PRESENTATION in terms of audibility, articulation and your ability to understand him or her:
6	Rate your professor's RESPONSES TO QUESTIONS:
7	Rate your professor's VISUAL PRESENTATIONS (blackboard, slides, overheads) for organization and legibility:
8	Rate the overall QUALITY OF TEACHING for this course:
9	Rate your professor's ATTITUDE TOWARD TEACHING this course:
10	Describe the AVAILABILITY of your professor outside of class. (Do not answer if you did not seriously seek help):
11	Would you like to nominate this instructor for an excellence in teaching award? If so, please include an additional comment below explaining why. If not, leave this question blank.
12	Your professor would like to know if there are some things you believe were done especially well in the teaching of this course.
13	Your professor would like to know what specific things you believe might be done to improve teaching of this course.