



MEMORANDUM

To: Executive Committee of Faculty Council (February 2, 2021)
Faculty Council (February 24, 2021)

From: Professor Ken Tallman
Chair, Teaching Methods and Resources Committee

Date: January 19, 2021; Revised February 24, 2021

Re: **Proposed Change to Course Evaluation Divisional Item 7**

REPORT CLASSIFICATION

This is a major policy matter that will be considered by the Executive Committee for endorsing and forwarding to Faculty Council for vote as a regular motion (requiring a simple majority of members present and voting to carry).

PROCESS AND CONSULTATION

TMRC is comprised of undergraduate and graduate students, faculty representatives from each engineering undergraduate program, Engineering Science, ISTEP, and an alumni member. As well, TMRC has representation from the Registrar's Office, the Engineering Education Technology Office, the Engineering and Computer Science Library, the Vice-Dean, Undergraduate, and the Vice-Dean, First Year.

With the goal of improving the student evaluation of teaching (SET) at the Faculty of Applied Science and Engineering, the Teaching Methods and Resources Committee (TMRC) has reviewed the FASE Divisional Items in the Course Evaluations, including Divisional Item 7.

In forming a motion to remove Divisional Item 7 from the course evaluations, TMRC consulted the Centre for Teaching Support & Innovation (CTSI) and the Engineering Society. Additionally, TMRC conducted a survey of FASE faculty and students; conducted a student focus group; analyzed other U of T Faculties' Course Evaluation Items; and surveyed literature studying course evaluations.

Based on discussion as well as the consultation and research described in Appendix 1, the TMRC members were divided in their support of the motion to remove Divisional Item 7. The committee agreed, however, that Faculty Council should have the opportunity to vote on this important matter.

MOTION FOR FACULTY COUNCIL

THAT the Faculty's Divisional Item 7, "What is your overall rating of the instructor as a teacher?", as described in Appendix 1 to Report 3680 Revised, be removed from the course evaluations, effective February 24, 2021.

Faculty of Applied Science and Engineering**Teaching Methods and Resources Committee*****Proposal for Revising Divisional Items in Course Evaluations******January 19, 2021*****1.0 MOTIVATION**

In Fall 2019, The Teaching Methods and Resources Committee (TMRC) voted to review the FASE Divisional Items in the Course Evaluations. The committee decided on this review for three reasons: 1) over five years had passed since the Divisional Items had been introduced and a review was now due, 2) some items (questions) appeared to provide little guidance for instructors seeking relevant and actionable feedback or for administrators seeking useful data, and 3) some items appeared to overlap with others, creating unwanted redundancy in an already demanding evaluation form.

2.0 PURPOSE OF COURSE EVALUATIONS

According to [1], student evaluation of teaching (SET) serves three purposes: “(a) improving teaching quality, (b) providing input for appraisal exercises (e.g., tenure/promotion decisions), and (c) providing evidence for institutional accountability (e.g., demonstrating the presence of adequate procedures for ensuring teaching quality).” Thus, as SET serves multiple stakeholders, some evaluation items will better meet the needs of different stakeholders. For instance, instructors, wanting formative feedback, will tend to favour “multidimensional” questions that address particular aspects of their teaching and their courses, e.g., the quality of course organization and feedback to students. Administrators and institutions, on the other hand, will use SET results for more summative purposes, looking for “unidimensional” and “global” items that measure overall performance [2].

3.0 INSTITUTIONAL ITEMS

There are 8 Institutional Items, asked of all U of T students. In addition to the 8 items, Items 1-5 are calculated into a multidimensional summary score, the Institutional Composite Mean (ICM), complementing the unidimensional global Item 6. Unlike the Divisional Items, the Institutional Items include two open-ended questions, allowing students to provide detailed feedback.

1. I found the course intellectually stimulating.
2. The course provided me with a deeper understanding of the subject matter.
3. The instructor created a course atmosphere that was conducive to my learning.
4. Course projects, assignments, tests and/or exams improved my understanding of the course material.

5. Course projects, assignments, tests and/or exams provided opportunity for me to demonstrate an understanding of the course material.
6. Overall, the quality of my learning experience in this course was ...
7. Please comment on the overall quality of the instruction in this course. [open-ended]
8. Please comment on any assistance that was available to support your learning in this course. [open-ended]

Though the Institutional Items are set by the University, and are thus outside the TMRC review, the TMRC recognizes that the Divisional Items must work in harmony with the Institutional Items.

4.0 DIVISIONAL ITEMS

The focus for the TMRC review were these Divisional Items.

1. The course helped me improve my ability to formulate, analyze and solve problems.
2. The instructor related course concepts to practical applications and/or current research.
3. Compared to other courses, the workload for this course was (*very light, light, average, heavy, very heavy*).
4. The instructor explained how the course concepts related to other courses.
5. The feedback I received on tests, assignments, labs, and/or projects provided guidance on how to improve my understanding of course materials.
6. The instructor explained what students are expected to learn in the course.
7. What is your overall rating of the instructor as a teacher?
8. The course instructor delivered the course material in a clear and organized manner.

The Centre for Teaching Support and Innovation (CTSI) provided TMRC the following guidelines for shaping our recommendations: TMRC could recommend to add one item to the existing Divisional Items without removing another; TMRC could recommend to remove any items without necessarily replacing them; and TMRC could recommend replacement items drawn from the existing U of T Course Evaluation Item Bank or recommend items newly created or modified by the TMRC in consultation with CTSI.

5.0 RATIONALE FOR TMRC RECOMMENDATION

The motion to remove Divisional 7 came with considerable debate, and, as noted, the TMRC members remained divided in their support of the motion. The TMRC agreed, however, that Faculty Council should have the opportunity to vote on this important item, giving all Faculty Council members the opportunity to have their voices heard.

Below are key points in support of Divisional 7:

1. It provides an “overall” assessment of the course instructor.
2. In the TMRC Survey of the Divisional Items, January 16-23, 2020, completed by 77 FASE faculty members and 245 FASE students, over 44% of our faculty and over 54% of our

students gave Div. 7 the highest score, 10/10, when asked “How appropriate do you feel the question is for use in course evaluations for all engineering courses?” (Appendix A).

3. It is an instructor-based item in a survey dominated by course-based items.
4. It allows students to rank instructors without considering externally imposed criteria.
5. It is concise and direct.
6. It has a history: people are familiar with it, and for many instructors and administrators, it is the first item they consider when viewing evaluation results.

Below are key points opposed to Divisional 7:

1. It provides no clear criteria for forming a judgment, making it unclear what students are assessing (Appendix C).
2. Global items about instructors such as Div. 7 are less reliable than specific items and are thought by some to be more prone to extraneous effect [3]. Extraneous effects include class size, course year, course type, instructor rank, and bias such as instructor gender [1] or race [5].
3. As an overall global rating, it is highly correlated with Ins. 3 and Ins. 6, and Ins. 6 is highly correlated to the Institutional Composite Mean (ICM), which means Div. 7 is also highly correlated to the ICM (Appendix D).
4. The ICM and Ins. 6 are perhaps better global indicators than Div. 7, as both focus on the course rather than the instructor, minimizing the risk of bias.
5. On the advice of the CTSI and the Vice-Provost’s Office, no other Faculties at the University of Toronto include this question in their Divisional Items.

The literature discussing whether there is student bias in course evaluations is inconclusive. A paper published in 1993 observes that “a majority of studies have found that male and female college teachers do not differ in the global ratings they receive from students,” noting that students tend to give slightly higher ratings to same-gendered than to opposite-gendered teachers [6]. However, an earlier 1987 study of 1,000 male and female college students found that male students gave female professors “significantly poorer” ratings than they gave male professors. The authors concluded, however, that although the differences found in the study were statistically significant, the “magnitude of the mean differences actually was quite small” [7]. A more recent 2000 study, which included over 700 classes across a variety of academic disciplines from a number of two-year and four-year colleges and universities, presented this concluding remark: “Is there Gender Bias in Student Evaluations of Teaching? The results reflect some same gender preferences, particularly in female students rating female teachers. But the difference in ratings, though statistically significant, are not large and should not make much difference in personnel decisions” [8]. This study notes that although “research has found potentially biasing factors to be of little or no influence, bias studies continue to play a central role in the recent literature” [8].

6.0 CONCLUSION

TMRC is divided in its support of the motion to remove Divisional Item 7 “What is your overall rating of the instructor as a teacher?” from the course evaluations. The TMRC agrees, however, that the Faculty Council should have the opportunity to vote on this important matter and decide for itself whether to remove Divisional 7.

LIST OF REFERENCES

1. Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research, 83* (4), 598-642.
2. Abrami, P. (1989). How should we use student ratings to evaluate teaching? *Research in Higher Education, 30* (2), 221-227.
3. Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist, 52* (11), 1187-1197.
4. Frey, P. W. (1978). A two-dimensional analysis of student ratings of instruction. *Research in Higher Education, 9* (1), 69-91.
5. Smith, B. P., & Hawkins, B. (2011). Examining student evaluations of Black college faculty: Does race matter? *The Journal of Negro Education, 149-162*.
6. Feldman, K. (1993). Views of male and female college teachers: Part II: Evidence from students' evaluations of their classroom teachers. *Research in Higher Education, 34* (2), 151-211.
7. Basow, S.A., & Silberg, N.T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology, 79* (3), 308-314.
8. Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching. *The Journal of Higher Education, 71*(1), 17-33.

OTHER SOURCES CONSULTED

Hativa, N. (2014). *Student ratings of instruction: recognizing effective teaching*. Oron Publications: USA.

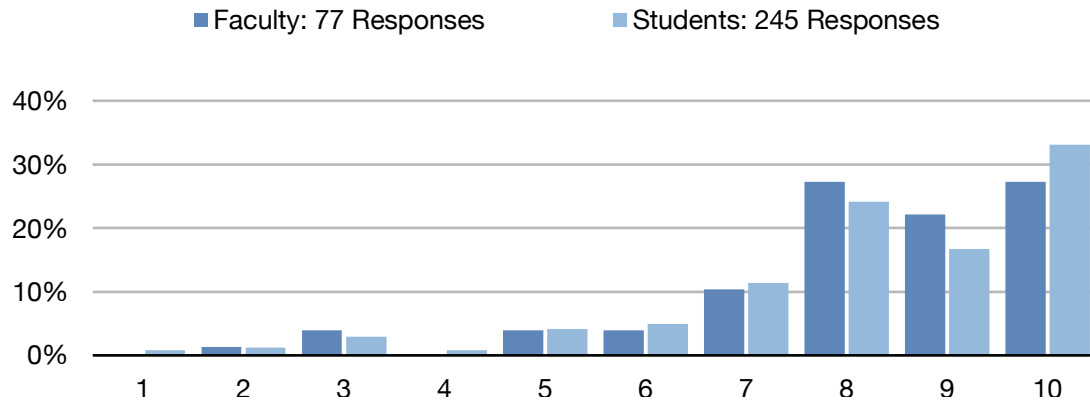
Marsh, H. W. (2007)b. Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. *In The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-383). Springer, Dordrecht.

Miller, J., & Chamberlin, M. (2000). Women are teachers, men are professors: A study of student perceptions. *Teaching Sociology, 28* (4), 283-298.

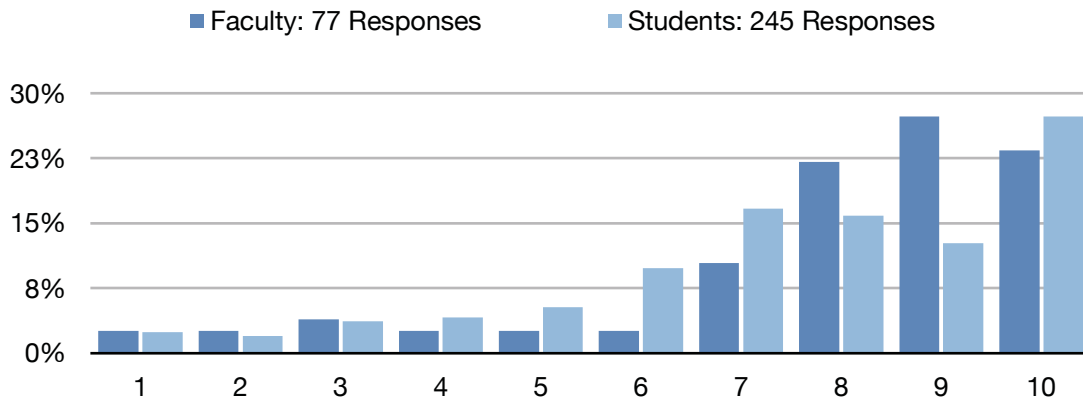
Ryalls, K., Benton, S., Barr, J., & Li, D. (2016). Response to "bias against female instructors". IDEA Research and Papers, Editorial Notes.

Appendix A: TMRC Survey of Divisional Items, January 16-23 2020 Results

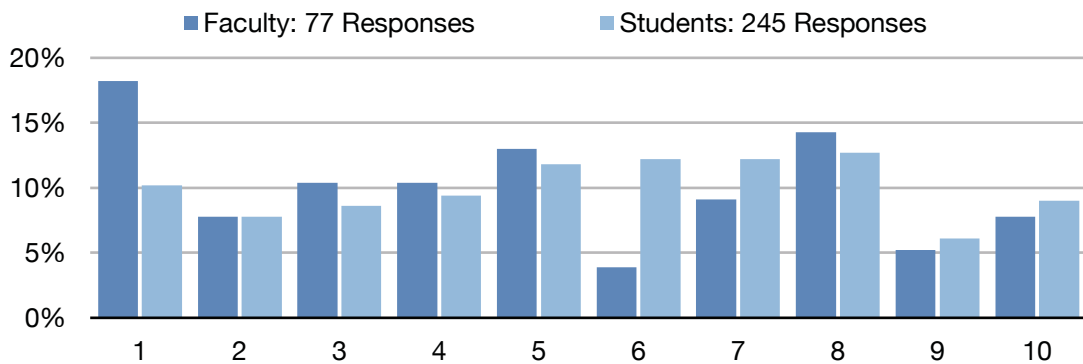
Q1: The course helped me to improve my ability to formulate, analyze and solve problems.



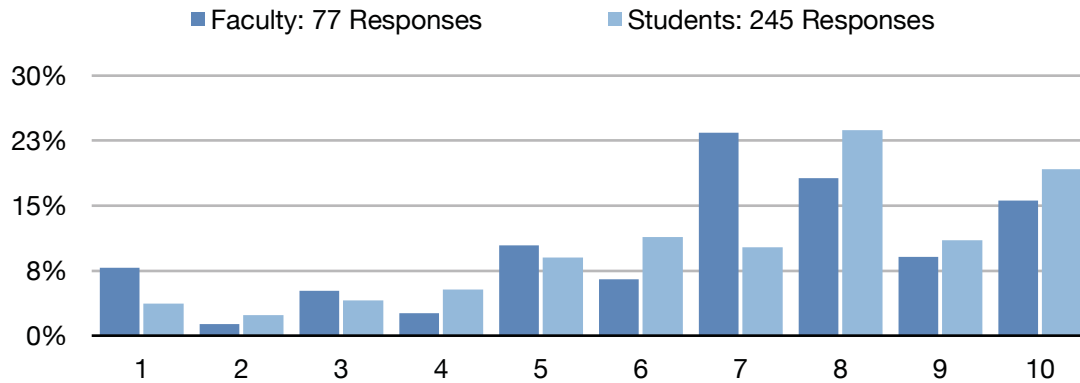
Q2: The instructor related course concepts to practical applications and/or current research.



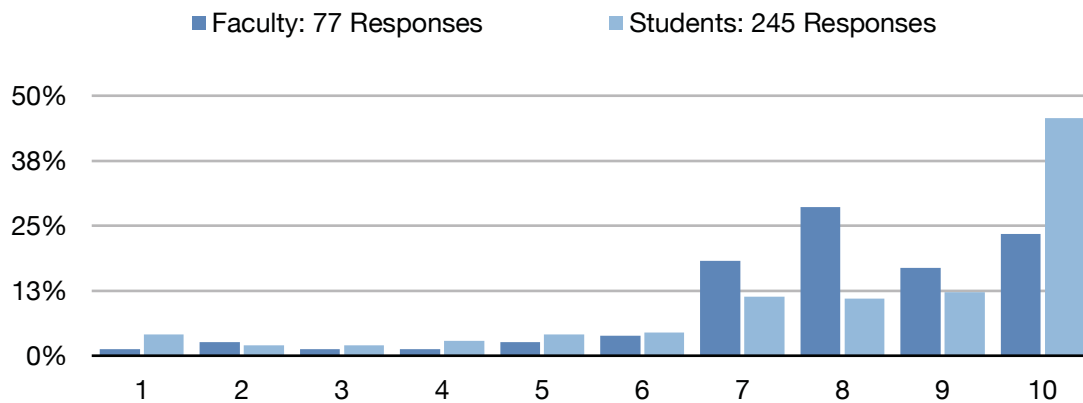
Q3: The course expanded my understanding of the ethical and environmental concerning engineering in society.



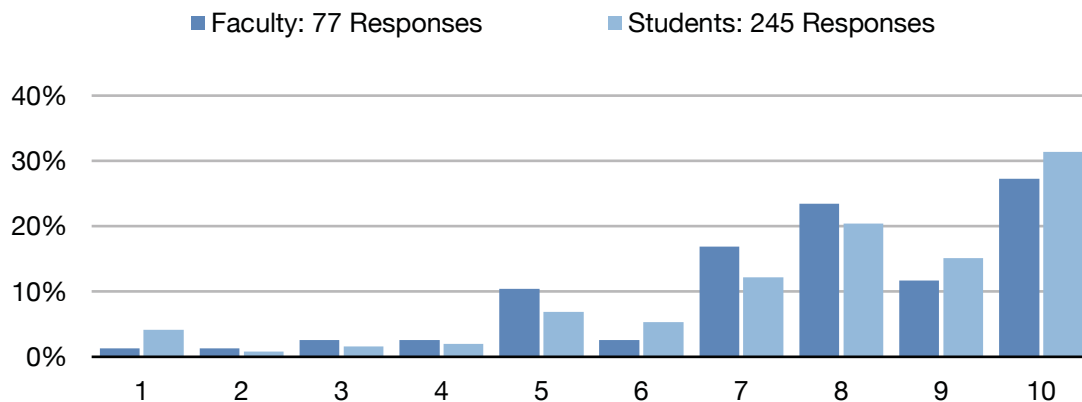
Q4: The instructor explained how the course concepts related to other courses.



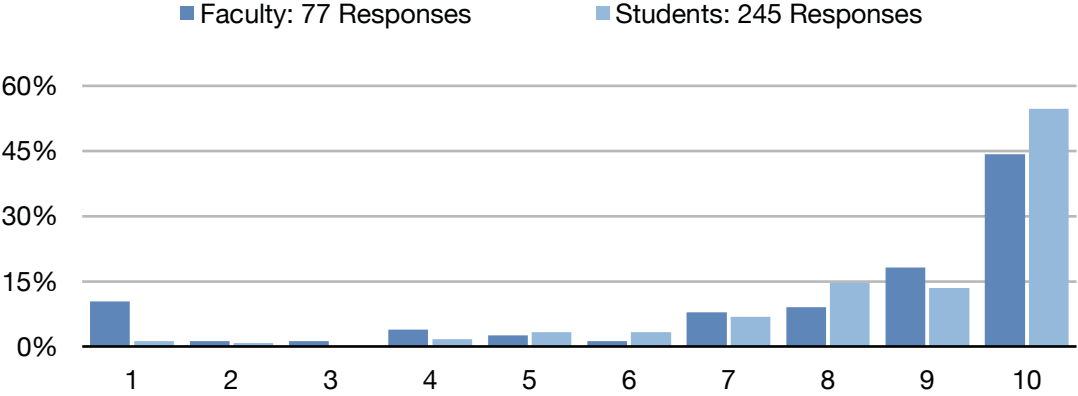
Q5: The feedback I received on tests, assignments, labs, and/or projects provided guidance to improve my understanding of course materials.



Q6: The instructor explained what students are expected to learn in the course.



Q7: What is your overall rating of the instructor as a teacher?



Appendix B: TMRC Survey Faculty and Student Comments

As part of the January 16-23 survey, we asked faculty and students two open-ended questions. Below is a selection of representative responses.

1. "Please comment on the reasoning behind any of the ratings you provided."

Student Responses:

- "The course expanded my understanding of the ethical and environmental issues concerning engineering in society." - I gave 2/10, because it could be useful for some courses which have intensive practical aspects (such as ESP, Design project), but not relevant for the courses which are teaching theories (such as programming fundamentals).
- Couldn't really relate most course material to environmental/social issues, maybe because we study BAsC.
- Ethics and environmental related issues are not relevant for all courses (i.e. differential equations and other math heavy courses), only actually useful for project-based courses.
- Honestly, if an instructor doesn't explicitly state in the syllabus why this course matters for other courses then I don't care if it somewhat semi-relates down the line.
- I don't think it is fair for math professors to be expected to tie their work to ethics/ the environment. I think this is fair for design courses etc., but for some of the theoretical math courses, it is not possible given the amount of content needing to be taught in a short time.
- I feel that the elements of the way a course is structured and planned out should be kept different from the way the professors teach and guide the students while collecting the feedback. I feel that they are two very separate things.
- I think it's super important to link the concepts we are learning in the classroom with real-life applications/research that is currently going on. At the same time, I'm not sure how linking classes together has a positive impact on what the professor is teaching - mentioning which concepts are taken from which classes might be helpful, or which future classes would build on the topics that are currently being taught, but I'm not sure we need such a big focus on this. I think asking questions about the instructor is very important, but I think they could be phrased better. Also, I feel as if the ethics/environmental question is not relevant to all the courses and shouldn't be mandatory for all of them.
- Not every course will aid in problem solving (ie APS100), and many of the courses don't significantly increase problem solving ability. However, feedback and expectations are integral to growth: if you aren't given correct feedback you can never improve, thus having courses with good feedback is essential. Similarly, a lack of clear expectations makes starting a problem very difficult: much more than what it needs to be.
- Overall rating could potentially just be biased towards one's personal opinion of the prof, instead overall rating should just be calculated (e.g. average or weighted average) of the other responses.

- Professors seem to rarely actually connect their courses to other courses, and I'm not sure how important that even is to determining the quality of the course. Also, we are rarely given much feedback in any course assignments/exams...
- Talk about real world experience. Talk about being able to explain things concisely. Talk about teaching ability. I honestly don't care how this course applies to other courses and I really don't see why that would apply to anything
- The course evaluations need to be less wordy. I understand courses can connect to research, other courses, and ethics, but most courses simply don't and that's okay because they don't need to. To me those questions are meaningless.

Also, the open-ended questions are worded very poorly and I'm always confused what I'm expected to write in them

- Think the most important feedback for courses should be focused on how well the course is able to set you up in that area. For example, after taking this course, how confident are you in your ability to apply these concepts to the real world?

Faculty Responses:

- All of the questions that are written in a way that allows for targeting the instructor (rather than course design, course delivery, etc) allow for unconscious bias to creep in. It would really help out female and URM instructors if the questions were written in a way that excluded the term 'instructor' but still captured the concept of the question. This is pretty straight forward to do for all of the questions listed that have the word 'instructor' in them.
- Ethics and environment are grouped - students in my classes don't differentiate between the two, as far as I can tell. The resulting question lacks clarity, and students (at least in my experience) give very low scores - which is fine, it just doesn't help me to know if i'm properly addressing ethics (because talking about the environment doesn't really make sense for my courses).
- Ethics are not part of any of my courses.
- Usually it's on the first day of the semester that the course's relevance and its connection to other courses were communicated to students. Those kind of references will only be made sparsely throughout the term when appropriate. So when you ask those kind of question at the end of the semester, most students would have little or nothing to say about it.
- I do not feel that comparisons to other courses should be assessed, unless the course under consideration is a prerequisite for another course. Also, not all courses have an environmental/ethical component; this question should be included only for pertinent courses
- I think all of the questions are well chosen and specific enough to elicit actionable feedback EXCEPT the poorly designed question on overall quality of the instructor. This item was retained from the previous course evaluation scheme primarily because Department Chairs said they needed to have continuity of information to evaluate instructors. We now have sufficient data with the current system to ensure that continuity of instructor evaluation

- I think all questions are very much on track and issues we should be concerned with. I know the final question is controversial to the purists but I have not been persuaded it is inappropriate
- Question 3 is not applicable to some courses such as mathematics, physics, etc. Question 4 is not applicable in many cases when concepts taught are standalone.
- Questions related to clarity/meeting objectives of the course are very relevant, as are those relating course material to the larger curriculum; questions getting into more specific details about course material (which assume homogeneity across all engineering courses) less so. The question rating the professor is inappropriate; studies have shown that the question says less about teaching effectiveness than about students' bias based on gender, race, etc.
- Rating of professors rather than learning is full of biases. The question on problem solving promotes rote learning. Questions around designing or creating would be more appropriate
- The "overall rating" of the instructor is much too vague for students - this question can easily become a popularity contest, e.g., the instructed material was hard, therefore we hate the instructor...
- The final question is addressed by the University questions - and could be eliminated.
- I argue that many students will perceive the number of questions as excessive, and answer them superficially, reducing the quality of the feedback and ratings we receive. I believe that 5-8 well-formulated questions (in total) would be sufficient and gather more effective feedback.
- The overall teacher question is not validated and has too many facets. Students cannot give a good answer to that as they can confuse likeability with teaching ability and other such confounding factors.
- There is a lot of potential for bias especially towards instructors who identify as female and/or are visible minorities. I feel strongly that the last question which rates the instructor should be dropped.

Question 2: "Are there other aspects of teaching and learning that you would like to include in the course evaluation?"

Student Responses:

- A question asking about the atmosphere the professor created in class, and if they created opportunities for the class to interact or openly ask questions.
- Amount of new material that you learned.
- Relating course concepts to job experience/ PEY experience.
- Can this course be improved? what else can we incorporate or remove into the syllabus that would improve the learning
- clarity and organization
- Does the professor care about students' learning in their class?
- The marks I received in the course felt fair and justified.
- The instructor was passionate about the course.
- How clear is your instructor when explaining ideas etc?

- How relevant the course is to your chosen discipline/career path rather than just how intellectually stimulating it is
- The school must do more to employ professors who actually want to teach.
- some instructors may have good lectures, but the tests and exams reward memorization over understanding concepts.
- I think there should be something about how well the course was organized.
- I would like the questions to be more actionable and have takeaways that the profs can improve on.
- Instructor should try explaining every concept in layman terms so that a non-expert can understand.
- The "quality of instruction" question is vague and confusing. I never know what exactly the "quality of instruction" means and usually use that section rant about the course.
- Is lecture content introduced in a structured, well-paced way? Is there continuity between lectures? Are lectures clear and understandable?
- Sometimes teachers are nice and support learning but they can't teach. Or they just don't teach in an organized direct way.
- Timeliness of marking or responding to questions.

Faculty Responses:

- Heavy course load, which previous generations accepted and leveraged to advance their learning, is now a lethal weapon used against instructors. Perhaps the solution is not to alter the evaluation criteria but to admit undergrad students more selectively.
- Did you feel that the course was run in a clear and efficient manner?
- How high was the workload of this course as compared to other courses you have taken in the same semester?
- I would like the course reviews to serious take into account bias (towards gender, skin colour and accent) in its design.
- Institutional questions 6 and 7 and divisional question 7 are very close and highly correlated (I'd assume, again, since the wording is very close). I would prefer if the divisional questions actually asked more granular things - "The pacing of the course was suitable and allowed students with different levels of knowledge to succeed." Or "the instructor effectively used lecture time..". Or "Lecture materials were appropriate for the material." Some of these are included in other sections or can be selected. Help me to determine which parts I'm getting right and which parts need improvements or changes.
- It would be good to add some information to the evaluation reports sent back to instructors on the impact of class size on average scores. From the course evaluation analysis completed last year by the university, it was clear that course size has a significant impact on average course evaluation scores (0.5/5 change from small courses to large courses). Including this information on the course evaluation report would be helpful. I teach some small and some large courses, so adding this information would help me interpret my scores better.

- No. Not because there aren't other aspects worth exploring, but because we need to resist the urge to ask more questions of our students. Like this survey, course evaluations should be quick and allow for freeform expansion.
- There seems a need to find out the student's workload in a course.
- There should be less questions, not more. Students can elaborate in the free-form fields on specific aspects of teaching and learning.
- Ways to help instructors provide actionable feedback.

Appendix C: TMRC Focus Group, Winter 2019

The TMRC spoke with class representatives across the Faculty, with students across all years of Engineering Science, and with members of the Engineering Society. TMRC asked the following question: “How do students understand Ins.3 (“The instructor created a course atmosphere that was conducive to my learning”) and Div. 7 (“What is your overall rating of the instructor as a teacher?”)? As Table 1 shows, students understand these questions in very different ways.

Table 1: Students Understanding of Ins. 3 and Div. 7

Ins.3: The instructor created a course atmosphere that was conducive to my learning	Div. 7 What is your overall rating of the instructor as a teacher?
<ul style="list-style-type: none"> • Accessibility of instructor • Instructor accountability • Encouragement of student learning • Does the instructor want to connect with students • Classroom management • Can I easily take notes in the class? • Will I ask questions in the class? • How do I feel in the class? • How does my learning style fit? 	<ul style="list-style-type: none"> • Clarity and organization • Speaking ability • Use of examples and tips • Lecture style (e.g., enthusiasm) • Punctuality • Engagement of the class during instruction • Use of teaching aids • Creation of supplemental materials • Conceptual understanding • My final grade/my understanding • Instructor as a person (e.g., likability) • Personal opinion about the instructor • “Gut impression” of the instructor

Appendix D: CTSI Correlations Between Institutional and Divisional Items.

As table 1 shows, there are notably high correlations between Div.7 and Ins.3 and Ins.6.

Table 1: Correlations Between Institutional and Divisional Items

Item	Ins1	Ins2	Ins3	Ins4	Ins5	Ins6	Div1	Div2	Div3	Div4	Div5	Div6	Div7
Ins1	1												
Ins2	0.78	1											
Ins3	0.62	0.67	1										
Ins4	0.62	0.67	0.59	1									
Ins5	0.60	0.65	0.58	0.82	1								
Ins6	0.71	0.75	0.78	0.71	0.71	1							
Div1	0.72	0.74	0.63	0.69	0.68	0.75	1						
Div2	0.52	0.57	0.66	0.5	0.5	0.62	0.57	1					
Div3	0.37	0.39	0.44	0.4	0.41	0.48	0.44	0.54	1				
Div4	0.52	0.54	0.65	0.5	0.5	0.63	0.6	0.68	0.57	1			
Div5	0.53	0.55	0.55	0.64	0.65	0.65	0.63	0.49	0.5	0.56	1		
Div6	0.56	0.62	0.74	0.59	0.6	0.71	0.63	0.64	0.43	0.65	0.6	1	
Div7	0.62	0.67	0.88	0.6	0.59	0.81	0.66	0.68	0.46	0.68	0.58	0.78	1