

Expert Supplemental Report on
Student Evaluations of Teaching (SET)
Richard L. Freishtat, Ph.D.

12 December 2016

Prepared for Ryerson Faculty Association (“RFA”)

I. **BACKGROUND**

- A. After providing an expert report on Student Evaluations of Teaching (SETs) on September 30, 2016, I have been asked by the Ryerson Faculty Association to provide an additional Supplementary Report assessing the specific SET tool used at Ryerson University. This tool, referred to as "Faculty Course Surveys" (FCS), is found in Appendix F of the collective agreement between the University and the Faculty Association.

II. **SUMMARY OF OPINIONS**

A. Appropriateness of Faculty Course Survey (FCS) questions

1. Students are not in a position to evaluate or comment upon the effectiveness of a course or instructor.
2. As explained in section B #2 of my expert report:
 - a) Student Evaluations of Teaching (SETs) are primarily measures of student satisfaction with their experience in a course.
 - b) A student's satisfaction with their experience in a course depends on many confounding factors that have nothing to do with the instructor's teaching effectiveness. Many of the factors that affect SETs are not what should be affecting SETs (e.g., Was the student earning the grades they thought they deserved throughout the course? Was this a required course on a topic the student did not wish to take? Did the student find the instructor's accent or appearance to be pleasant or unpleasant?).
 - c) SETs do not accurately measure a faculty member's teaching effectiveness. There is no compelling correlation between student learning and more highly rated instructors.
3. As explained in section B #5 of my expert report: "Current students are well positioned to comment on their own experience of the class and inputs like: instructor's ability to communicate clearly, enjoyment, difficulty or ease, engagement or boredom, if an elective then whether they plan to take a sequel course, favorite/least favorite part of a course, whether they would recommend this course to other students, hours spent per week outside of class, and background information about pre-requisites taken or other courses in the field."

4. Attention should still be paid to the fact that “even within these areas of student ratings, student comments can be valuable, but interpreting them is troublesome.”

5. Appropriately focused questions contained in the FCS to provide information about a student’s experience and background information include none from the University Standard Questions, and six from the Optional Course Specific Questions:
 - a) Optional Course Specific Question, Laboratory Courses Question #3 “I feel free to ask for assistance and to ask questions.”
 - b) Optional Course Specific Question, Laboratory Courses Question #8 “Lab assignments are interesting and stimulating.”
 - c) Optional Course Specific Question, Discussion/Tutorials/Seminars Question #4 “Discussion in this course is stimulating.”
 - d) Optional Course Specific Question, Discussion/Tutorials/Seminars Question #7 “I feel encouraged to participate in the discussions.”
 - e) Optional Course Specific Question, Clinical/Field Placements Question #1 “Prior coursework prepared me to handle the clinical tasks.”
 - f) Optional Course Specific Question, Other Course Elements Question #2 “I am evaluated for my individual contribution to group work in this course.”
 - g) *Note that while these questions are focused on areas of classroom experience students are able to comment upon, ratings themselves are still not measures in evaluating teaching effectiveness, nor are they formative in nature without accompanying qualitative questions seeking comments that inform why a student experienced the course in a given way.

6. Inappropriately focused questions include all others in the FCS, and as such should not be asked at all, or in current form. Here I offer a detailed breakdown of the University Standard Questions and the areas of foci they inappropriately address or invite bias. Every Optional Course Specific Question falls into one or more of these same categories - other than those few mentioned specifically in the section above.
 - a) As explained in section B #3 of my expert report: “Students should not be used to rate the adequacy, relevance, and timeliness of the course content nor the breadth of the instructor’s knowledge and

scholarship.”

(1) University Standard Question #1 “The instructor is knowledgeable about the course material.”

b) As explained in section B #3 of my expert report: “Most students lack the expertise needed to comment on whether the teaching methods used were appropriate for the course.” And, as explained in section D #2 of my expert report: “SET are affected by gender biases and stereotypes ... and result in lower ratings on every aspect of teaching, including putatively objective measures.”

(1) University Standard Question #7 “The course handouts/postings contain all of the information I need about the organization and operation of this course.”

(a) The choice of handouts/postings, and their organization into the operation of a course is a part of the teaching method used by an instructor. It is a choice of content delivery, much like the choice of lecture versus discussion. Teaching methods of all kinds and forms matter, and most students do not possess the expertise necessary to comment on it.

(2) University Standard Question #11 “The course is well organized and managed.”

(a) Whether a course is organized and managed well fall under the purview of teaching methods. Different pedagogies drive organization and management of a class in uniquely different and varied ways. How an instructor organizes and manages a course relying heavily on Project-Based Learning would be different than a course relying solely on lecture. Most students do not have the expertise to comment upon this, and many other, necessary distinctions.

(3) University Standard Question #14 “The way this course is taught helps me to learn.”

(a) Most students lack the expertise to comment upon their learning in a course, and in fact most confound their relative experience with learning. There is no compelling correlation between student learning and more highly rated instructors.

- (b) Most students lack the expertise to comment on any aspect of a course or instructor, meaning in addition to the topic of learning, the phrasing of “this course” and “is taught” are areas students should not be asked to comment upon.

- c) As explained in section B #3 of my expert report: “Most students lack the expertise needed to comment on ... if the content covered was appropriate for the course.”
 - (1) University Standard Question #4 “Concepts are clearly explained with appropriate use of examples.”
 - (a) Most students lack the expertise to comment on any aspect of a course or instructor. The phrasing of “Concepts are clearly explained...” asks students to do exactly this - to comment upon how the instructor explains concepts, instead of how they individually experienced concept explanations.
 - (b) The choice and use of examples is part of the content covered in a course. Content is not restricted the concepts themselves. It is and should include the breadth and depth of how those concepts are covered - meaning that examples provided to illustrate those concepts are course content. Most students lack the expertise to comment on the appropriateness of course content, and therefore the use of examples as well.

- d) As explained in section B #3 of my expert report: “Most students lack the expertise needed to comment on ... if methods of student engagement used were appropriate to the level and content of the course.”
 - (1) University Standard Question #3 “The instructor stimulates my interest in the subject.”
 - (a) This question in particular is poorly worded, since students can comment upon their engagement or boredom - but not the cause of it. It asks the student to comment on the instructor’s impact on their interest, which invites bias and confounding factors. If it asked about the student’s interest alone, it

would be more appropriate.

- e) As explained in section B #3 of my expert report: “Most students lack the expertise needed to comment on ... if the assignments were appropriate for promoting and assessing their own student learning.”
 - (1) University Standard Question #8 “The assessment methods, including tests, provide a fair evaluation of my learning.”

- f) As explained in section B #3 of my expert report: “Most students lack the expertise needed to comment on ... if what they learned has real world application, [and] if what they learned will help them in future classes.”
 - (1) University Standard Question #13 “This course provides a valuable learning experience.”

- g) As explained in section B #3 of my expert report: “Most students lack the expertise needed to comment on ... if the type of assistance, help or support given to students was appropriate to the learning goals of the class.”
 - (1) University Standard Question #6 “I get constructive feedback on my assignments.”
 - (a) While the phrasing appears to be focused on the student experience, the area of question focus is inappropriate. The question asks for a judgment about the helpfulness of instructor feedback on assignments, with the implication that it is helpful to the learning goals of the class - of which assignments should be directed. However, most students lack the expertise to comment on the type and quality of help in relation to the learning goals; in this case, help in the form of constructive feedback.

- h) As explained in section B #3 of my expert report: “Most students lack the expertise needed to comment on ... if the course instructor was excellent, average or poor overall.”
 - (1) University Standard Question #15 “Overall the faculty member was effective.”

i) As explained in section D #1 of my expert report: “Personal traits [and bias] that are considered strongly related to SETs have been shown on items related to instructor fairness [and] professor attitude.”

(1) University Standard Question #9 “Students are treated with fairness and respect.”

j) As explained in section D #2 of my expert report: “SET are affected by gender biases and stereotypes ... and students rated a male instructor higher especially on ‘enthusiasm,’ ‘showed interest in the subject,’ and ‘using a meaningful voice tone.’”

(1) University Standard Question #2 “The course material is presented with enthusiasm.”

k) As explained in section D #2 of my expert report: “SET are affected by gender biases and stereotypes ... and result in lower ratings on every aspect of teaching, including putatively objective measures such as the timeliness with which instructors return assignments.”

(1) University Standard Question #5 “I get timely feedback on my assignments.”

(2) University Standard Question #10 “The class meets as scheduled and on time.”

(3) University Standard Question #12 “The instructor is available for consultation as specified on the course handouts/postings.”

7. The structure of the mandatory questions, and other optional, questions is not appropriate. Asking more general questions first (not to mention the flaws within the questions themselves) invites more bias and confounding factors into the rating of those items. While not possible to account for, or prevent bias on items like these, it can be better approached by designing SETs that focus on course-specific experience and inputs first, before moving to more generally applicable student experience and input questions. In this way, when students rate a general question, it is done so after first considering specific aspects of course experience.

8. Overall, the FCS is poorly constructed and invites bias in virtually every response and rating item. Note the difference between appropriate and inappropriate items. Appropriate items ask the student to comment on themselves, their background, and their experience only. Inappropriate items ask the student to comment on the course or instructor, and the impact on them/their learning. These types of questions invite the most bias, and include areas of teaching and learning that students do not have sufficient expertise to comment upon.

B. The 5-point rating scale

1. A 5-point Scale, or any scale, is not appropriate to utilize in an FCS. The number of scale items, and numbers/labels of the items themselves, are essentially meaningless. The numbers/labels, and the difference between numbers/labels, do not necessarily mean the same thing to different students in the same or different courses. They presume that a low rating and high rating can balance to make an average rating. For SETs, there is no reason any of those things should be true. Giving students more or less rating options does nothing to improve the accuracy or reliability of the meaning of those numbers/labels.
2. SET ratings are ordinal categorical variables: The ratings fall in categories that have a natural order, from worst (1) to best (5). But the numbers are labels, not values. Replacing the numbers with descriptions means no information is lost. In fact, if ratings are to be used, it would be advantageous for the FCS students' complete to have equivalent full descriptive labels and not basic numbers. The advantage of descriptive labels, like those reflected in the collective agreement between the University and the Faculty Association, makes the temptation to average scores and represent teaching by a single number challenging. Any use of numbers in the scale reproduces the problems that are to be avoided. Regardless, descriptive labels themselves do not mean anything more than they otherwise do as highly subjective and relative comments on student experience that cannot be compared easily for evaluation purposes, or bias accounted for in those ratings.
3. The rating categories used in the FCS are good if scales are used - although if used, they should be done with caution and in conjunction with

qualitative questions in informing an evaluation of teaching effectiveness.

- a) Any rating category should have a clear valuation attached to it so students are clear about where their rating falls. Those valuations should have a clear label about how evaluation will be interpreted, hence the use of descriptive labels.
- b) The rating categories in the collective agreement themselves are fine (the numbers used on actual FCS's are not appropriate), except for the label "neither agree nor disagree". There is a 5-point scale, with "neither agree nor disagree" as the #3 item. But, for a student to neither agree nor disagree could imply many possible factors, including that the question does not apply or they have no opinion. To enable this label to exist on the scale as #3 places its relative importance at the mid-level of perceived effectiveness or student experience. The "neither agree nor disagree" response more appropriately should fall outside of the scale completely so as not to be construed as a 3 out of 5 and a mediocre rating. It is, in fact, a non-rating.

C. FCSs for on-line courses

1. The current FCS would not be appropriate for use in an on-line course. In the same way separate course-specific questions are articulated as optional items, course-specific modality questions should be utilized in an on-line course. On-line pedagogy is nuanced and in many ways fundamentally different in its application and approach from in-class pedagogy. This should be accounted for in asking questions about student experience specific to the online nature of the course, and in a way that would inform future improvements to the course and its delivery.
2. For example, appropriate FCS items for an on-line course would include: "How many on-line course I have previously taken." "I have experience with the following on-line tools: discussion board, chat, etc. - and to what level of experience." "I was interested in the on-line discussions. What was interesting/not interested about them?" "I found the on-line tutorials and exercises engaging. What was engaging/boring about them?" "I had an easy time adjusting to the experience of an on-line course. Why, or why not." "I utilized course links to web-resources and materials. Why, or why not?"