

[Slide 1]

Welcome to today's webinar brought to you by the Canadian Sociological Association.

(Re)Prioritizing Pedagogic Feedback: Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

In recent years, mounting evidence of the dubious validity of Student Evaluations of Teaching (SETs) has generated a series of challenges to their use in decision-making on faculty careers across North America.

However, largely absent from the scholarly literature on SETs is what to do with the open-ended feedback students provide on SETs, which faculty are expected to utilize to improve their teaching.

There is sparse research into how these qualitative comments affect faculty mental health, wellbeing, professional self-esteem, and pedagogical practise. Yet, anecdotal accounts of unfair, hostile, even harassing comments are increasing with the transition to online delivery of the survey instruments. The potential for unevenness in psychological burden that this imposes raises issues of equity, given that identity-based biases in SETs scores are well documented.

[Slide 2]

We invite you to engage with our panelists. The 'Chat' function is meant for sharing comments which will only be visible to panelists. Please utilize the 'Q & A' function to submit your questions. You may upvote questions to be prioritized.

[Slide 3]

I would now like to introduce you to our moderator, Dr. Rochelle Côté.

Dr. Côté is the current Chair of the Canadian Sociological Association's Policy, Ethics and Professional Concerns Subcommittee and will be discussing the highlights of the report with Dr. Kowalchuk as well as facilitating engagement with the audience.

[Slide 4]

Good afternoon everyone from sunny and warm St. John's on the East Coast of Canada today and welcome to the CSA webinar series where we will be talking about the foundational report being released by the association that looks at the faculty experiences with qualitative comments on student evaluations of teaching.

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

I would like to start by acknowledging the Indigenous Nations on whose lands we are meeting today (from wherever it is you may be) to thank them for the opportunity to present our thoughts on their territories. We would also like to pay our respects to Elders of these lands, both past and present.

[Slide 5]

Our guest speaker today is Dr. Lisa Kowalchuk who is a professor of sociology at the University of Guelph. She spear-headed this project when she was in my role as chair of the CSA Policy, Ethics, and Professional Concerns subcommittee.

[Slide 6]

I would also like to acknowledge Dr. Kowalchuk's collaborators.

Rachel La Touche is an Assistant Professor, Teaching Stream in the Sociology Department at the University of Toronto.

Rochelle Wijesingha is currently a Ph.D candidate in Sociology at McMaster University and a Senior Research Associate at the Diversity Institute.

I think it is fair to say that both of these smart women were instrumental in seeing this report through to completion. I don't think Dr. Kowalchuk would argue with me on that at all.

[Slide 7]

We will now get started and I will turn the webinar over to Dr. Kowalchuk to report on this research.

[Slide 8]

Hello everyone!

To all who are joining this event live, and also to those who will view it later on, thanks for your attention to this topic. I'm really honored and pleased to be part of the CSA Webinar series.

Over the next 30 minutes or so, I will be presenting highlights of a study co-authored by Rachel La Touche, Rochelle Wijesingha, and myself, on Faculty Experiences with Qualitative Comments in SETs. It is a survey-based study, and it was funded by the CSA, with support also provided by OCUFA, the Ontario Confederation of University Faculty Associations.

Some of you will have read the full report which was made available on the CSA website a few weeks ago. But whether you have or not, the presentation is intended to set the stage

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

for some productive discussion of our results and their implications for better ways to obtain, purpose and utilize open-ended feedback comments from students.

My two co-investigators, Rachel and Rochelle could not join us due to work commitments. They each have important fortés that made the study possible so I'll do my best reflect what I learned from them in working on the project.

What I'm going to today is talk about why and how we did this study, the main highlights of the study, and what my co-authors and I think should be done in light of those findings and related scholarship.

[Slide 9]

Why/Impetus

What was the impetus for a cross-Canada study focusing specifically on the open-ended feedback component of SETs?

There are 2 main motivators of the study. The first concerns a relative silence in the literature about SETs. We knew that the recent scholarship was making an increasingly compelling indictment of the **numerical scores of SETs surveys** – partly on the grounds of their validity as a metric for summative evaluation of teaching.

For those less familiar with the terminology, “summative” refers to the use of the SETs results to determine one's overall teaching performance and to rank it relative to your colleagues. There's been a growing reliance on the scores from SETs for that purpose, by universities throughout North America and much of the Western world, since the 1970s. It was increasingly clear that SETs scores are a vastly inadequate indicator, some have called them a pseudo-indicator, of teaching quality.

Just as disturbingly, the latest studies were also showing quite definitively that SETs scores are biased against faculty who belong to equity-seeking groups. They are essentially a source of occupational discrimination against already disadvantaged groups, when they're used to determine promotion, renewal, hiring, and performance ranking.

A testament to how impactful and definitive these recent studies have been, was the 2018 resolution of a grievance brought by the Ryerson University Faculty Association, against Ryerson university, over the use of SETs. The arbitrator ruled that SETs scores should no longer be used for tenure and promotion decisions at that university.

But SETs surveys typically also include set of open-ended questions for comments on courses and teaching. Yet until very recently there was very little research into what was going on in the content of these qualitative comments and how they affect those receiving them -- even the Ryerson ruling doesn't really talk about the comments, or what should be done to improve how they're obtained and used.

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

The **comments** in SETs may **seem** to count less in the summative appraisals of faculty but they do get taken into account by evaluating committees and professors are expected to engage with them to improve their teaching.

Now, a study we cite in our report finds that professors tend NOT to talk much with each other about comments that are degrading and mean-spirited, out of a sense of embarrassment and shame. Nevertheless, we could see more and more anecdotal essays being published observing that unfair, hostile, and even harassing comments were creeping into SETs.

Given the unequivocal finding that SETs **scores** are biased, to us this meant that the content of the **comments** may also be treating equity-seeking groups worse than others. This would mean potentially an uneven burden in terms of faculty mental health, wellbeing, professional self-esteem, and pedagogical practise. It would also mean uneven penalties and rewards flowing to faculty when their SETs comments were being examined by committees for summative purposes.

Having said that, though, it is not acceptable that **anyone** be subject to harassing, hostile comments in their evaluations. That was the first main impetus.

[Slide 10]

A second impetus has to do with how well or poorly the comments in SETs are serving our pedagogy and by extension our students' learning.

A word to the undergraduate students who are viewing this and engaging with our study; we believe that you can only benefit from long overdue improvements to the way your qualitative feedback on university teaching is obtained and processed. That is because the system that's been in place for several decades for SETs in general, and for the commentary component of SETs in particular, undermines the quality of teaching in many ways.

Devising ways to obtain more meaningful feedback that includes constructive criticism will open the door to greater improvements in pedagogy. We believe that students who learn how to provide good-faith feedback on both strengths and weaknesses on the courses they're taking, will take those skills into their careers and their life journeys.

So now that I've told you **why** we undertook this study, I'll describe **how** we did it.

[Slide 11]

Methodology

We focused our study on universities across Canada. The survey was translated into French so that it could be responded to by francophones in Quebec or elsewhere. *Incidentally we had 10 respondents who completed the survey in French.*

In terms of discipline we confined it to departments of sociology or those which housed sociology. We created the sampling frame out of e-mail addresses listed on department websites. These lists included contingent faculty, though we know those lists were incomplete or not up to date.

Using Qualtrics, we emailed the survey to 1,179 instructors in August 2019 and kept it open for one month. The response rate was 26.4%. After deleting those who did not complete the whole survey, the final “n” was 288.

[Slide12]

Let me show you the sample characteristics

Appendix B

A couple of interesting things to note

- just over 18% of the sample were contingent faculty
 - that’s much lower than the figure reported by the CCPA -- Canadian Centre for Policy Alternatives -- in 2018, in a report called Contract U which finds that it’s 53%
- this under-representation in our sample undoubtedly has to do with the limitations using department websites for contacts
- almost 2/3 were in the 41 to 59 age range
- far more women responded than men: 57% of the sample
- and there were 15 non-binary gender respondents
- 15% of the sample were members of racialized groups
 - considering this in relation to the figure of 21% for university faculty overall, reported by the CAUT, Canadian Association of University Teachers in 2018, in a study called *Underrepresented and Underpaid*
 - 21% is also the figure for the labour force as a whole
- we had only 6 indigenous respondents; we aggregated them with other respondents in racialized groups

[Back to Slide 11]

Concerning questions in the survey most of them were closed-ended, with most of the core questions of the scaled type and there were 2 main open-ended questions.

One asked people to share their most memorable or impactful qualitative comments received from past SETs; 122 people answered that. The other asked respondents to share anything else about student evaluations of teaching that we had not asked about; 152 respondents responded. These responses were open coded using NVivo 12 Plus.

[Slide 13]

Our survey questions and our hypotheses correspond to four main overarching research questions. I'll address the first three of them one by one, and I'll address the 4th one, on identity characteristics, in connection with the other three.

[Slide 14]

To what extent are faculty in Canadian sociology departments receiving pedagogically irrelevant, hostile, harassing comments and even ones that make them feel unsafe?

Informing the concern about the **pedagogical relevance** of the comments is that the ostensible purpose of SETs is to signal to professors what they're doing well or not so well, and how they can improve.

In terms of literature, as I mentioned, there's little systematic study of SETs open-ended comments but recent scholarship on SETs **scores** shows that ratings are influenced by things like the weather on the day the survey is done, or the hand-outs of sweets shortly beforehand.

The compelling research I mentioned earlier on identity-based bias in SETs scores, especially along lines of gender and racialization, is also germane to this issue because with women and racial or ethnic minorities receiving **lower scores**, clearly there is something else being responded to, other than teaching quality.

[Slide 15]

The survey item we used to get at this, was a question asking the extent to which SETs comments that respondents receive were a reflection of things unrelated to pedagogy. The ones we asked about were: grade expectations, class size, instructor attractiveness, gender, being a member of a racialized group, age, and perceived language proficiency or accent.

[Slide 16]

Figure 6

An overwhelming majority report that yes, that their SETs comments reflect criteria that have nothing to do with pedagogy. Percentages range from 70s to 90s, depending on the specific non-pedagogical criterion. As you can see, these largely have to do with ascribed characteristics of the instructors themselves.

[Slide 17]

Figure 7

For most of these non-pedagogical criteria, the percentage is higher for women than men.

[Slide 18]

Figure 8

The percentages are higher for LGBTQ respondents.

The differences between racialized and non-racialized faculty were not large but let's keep in mind that the percentages are quite high across the board.

[Slide 19]

Our survey's open-ended items reveal that even the positive comments that instructors receive were also pedagogically irrelevant: they hovered around terms like nice, best, and favorite. They were laden with gender-bias, with terms like "nice" and "caring" vs. brilliant aligning with gender stereotypes. In this respect, the results show how the combination of **age** and gender can be the focus of discriminatory feedback that actually punishes women for being knowledgeable.

One respondent in her late 50s observed: *"It was worse when I was teaching courses that they didn't like, such as statistics and theory. Even when I was knowledgeable, one student felt that since I knew so much, that it was time for me to retire."*

[Slide 20]

Our open-ended responses also reflect discrimination toward both younger **and** older faculty depending on the category of assessment. These experiences were only reported by women, and were intertwined with sexist remarks.

An illustrative example is Jordan (45), a self-described "young'ish woman of colour," who regularly receives comments about her "wardrobe and personal style". This reveals liberties that students take to evaluate young women on their looks and presentation of self.

Whereas Claudette (70) and Gloria (51), report increasingly negative and ageist comments on SETs over their teaching careers. Gloria comments:

"I am still the same teacher I always was but as I have aged -- no chili peppers for me on ratemyprofessor.com anymore -- my evaluations have gone downhill. I have also noticed that the comments often reflect my 'failure' to be maternal towards my students, and that this expectation, and my failure to do so, has increased as I have aged."

[Slide 21]

Still on the first of the three overarching questions, which is really a compound question: It's already bad news that SETs comments are being experienced as largely unhelpful to improving one's pedagogy but what about the other part of this overarching question: comments that veer into **hostile, harassing, or abusive terrain?**

Informing this question was a growing number of anecdotal published accounts over the past decade or so, of a deterioration in the civility and constructiveness of SETs comments, and a turn toward troll-like content. These accounts suggest that this is a worsening problem, thanks largely to the digitization of the survey instrument, although other factors are argued to play a role.

I think it's because of these growing concerns that a large-scale, systematic study of the comments was done in Australia by the National Tertiary Education Union in 2017. This is the only study we know of its kind, that focuses on faculty's reports of the comments they receive. This study finds that 60% of respondents receive abusive comments, and that these focus on things like appearance, religion, culture, sexuality, disability, and perceived English language ability. They also found these were unequally distributed across particular identity groups, so they were much higher for women, for non-binary gendered and other sexual minorities, for those born outside Australia, and for those who are racialized.

[Slide 22]

Figure 5

The survey item **we** used for this was what you see here in Figure 5.

A majority of our respondents, 61% receive such comments. This is identical to what was found in the Australian survey. In terms of how these correlate with key identity traits, the percentages are higher for women than men, and higher for racialized respondents, higher for LGBTQ, and higher for the older respondents.

We don't have a graphic for this, but of those 61% of respondents who received hostile, abusive, or bullying comments, 27% indicated that comments have made them feel unsafe. This varied strikingly by whether someone is in a racialized group, with 44% of these respondents, vs. 23% of non-racialized respondents, being made to feel unsafe.

[Slide 23]

Next, our second major overarching question: **What is the impact on emotional wellbeing and morale?** Might the comments be so demoralizing that they make people feel like exiting the field?

In terms of what within the literature was driving this concern. Certainly, we simply had a hunch about this. There is very little literature on the mental health of university instructors.

In terms of more precise data, the Australian study that I mentioned did look into this issue. This bolstered our belief that this was appropriate to look at. They found that of the instructors who received abusive comments, 70% experienced negative emotions, and sizeable minorities experienced physical stress symptoms like disturbances in sleep and appetite and that the negative emotional impact of SETs comments cause a sizeable minority of faculty, about a third, to wish they could leave teaching.

[Slide 24]

Figure 3

So how did we try to get at this in our survey? One of the items was a general question about emotional impact, as you can see in figure 3. About one third overall said yes to this question and it's strikingly higher for women, and for racialized faculty which is what one would expect from respondents who are getting more of those abusive and bullying comments.

We also examined intersectionality here. Respondents who identified with two or more equity-seeking groups much more often said that qualitative comments have had a negative impact on their sense of well-being (41%). For those who identified with one group, it was 38% and for those who identified with none, 21%.

[Slide 25]

Figure 4

We also asked more specifically about anxiety. On this item, a much higher percentage of the overall sample answered affirmatively and it's considerably higher for women and for LGBTQ respondents also for younger respondents. Again, when we looked for intersectional effects, 67% of those who identified with two or more equity-seeking groups answered "yes". For those who identified with one designated group it was 65%, and 46% for those who don't belong to any equity seeking group.

We did also ask if people had ever considered leaving academe because of negative SETs comments. 16% answered affirmatively, and there were no meaningful differences among the socio-demographic sub-groups. But when we aggregate those who identified with two or more equity-seeking groups, the percentage rises to 26% and it rises to 33% for contingent faculty.

[Slide 26]

Figure 2

We also asked about the impact of SETs comments on morale related to teaching. Overall, only one-fifth to just under one-third of respondents indicated that qualitative comments on SETs **increased** their self-esteem or confidence and again, with no differences meaningful enough to report among the socio-demographic sub-groups.

[Slide 27]

Our third overarching question concerns the impact of SETs comments on pedagogy.

The scholarship on SETs scores' effect on pedagogy made us want to see whether there's a parallel in the comments. With the numerical ratings of SETs, there is compelling evidence that they've had a perverse effect on learning. We know that as universities have grown to rely on SETs scores since the 1970s there has been grade inflation.

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

As well, there are studies showing that SETs scores vary proportionately with the grades that students expect, based on grades already received. There are also studies showing instructors' awareness that making a course easier, or grading more leniently, leads to better SETs scores, and that they tailor courses accordingly.

So, we wanted to know whether faculty avoid innovation and make courses easier to avoid hostile feedback?

[Slide 28]

Figure 12

The percentage of our respondents who curb innovation and reduce difficulty levels was about 40% overall - less than the majority, but substantial nonetheless; it's not what you would expect to see, given the ostensible purpose of SETs comments. There were also noteworthy differences by demographic sub-group.

[Slide 29]

Figure 13

The percentages of women who avoided innovation, reduced the amount of material, and reduced difficulty level, were consistently higher than men.

The percentages of women who avoided innovation (46%), reduced the amount of material (51%), and reduced difficulty level (46%), were consistently higher than men; where it was 27%, 34%, and 34% respectively.

[Slide 30]

Figure 14

There's also a correlation with being a member of a racialized group: 49% of racialized respondents reduced the amount of material, and 46% reduced the difficulty level, compared to 40% and 39% of their non-racialized peers.

[Slide 31]

Figure 16

There is a difference by age. Considerably more of those aged 40 and under avoided innovation, reduced the amount of material, and reduced the difficulty level.

Among those aged 40 and under, **45%** avoided innovation (compared to 33% for those 60 and older), **51%** reduced amount of material (compared to 35%), and **49%** reduced difficulty level (vs. 37%).

[Slide 32]

On the question of the impact on pedagogy, there is an interesting contradiction in some of the results. Some of the items show that only a minority of the respondents see SETs as any kind of useful signpost about their teaching.

For example, only one-third of the overall sample believe that SETs comments are an effective means of determining teaching effectiveness and fully 38% believe that the comments have contributed to eroding academic standards. Yet, almost two-thirds reported that SETs comments are helpful in improving course design. We can only speculate on why this is the case. It might be that faculty are defining improvement as pre-empting abusive feedback to achieve better SETs in future courses –both in scores and comments.

[Slide 33]

In our qualitative results on this set of questions, we saw that respondents struggled to make sense of feedback from within the same course that was polarized and thus contradictory, with the same aspects of the course eliciting both highly negative and highly positive appraisal.

They also reported that innovations they invested greatly in, were rejected in the comments as having no pedagogic value. This reinforces the substantial minorities of respondents in the closed ended items who avoided innovation, reduced the difficulty level, and believed that the comments are eroding standards. Some reported outright dishonesty in the comments. Here is an illustrative remark along those lines: “I regularly have students lie -- saying that I only read verbatim from the slides, that I don't give them instructions.”

[Slide 34]

Contingent faculty

We wanted to capture the experiences that **contingent faculty** have with SETs. There is little systematic research on this but a number of anecdotal essays have been published that point to very disturbing uses and effects of SETs both scores and comments, on contingent faculty. I do want to flag a new in-depth, forthcoming study of this, the only one we know of, by Smele et al.

In our study 52 of our 288 respondents were contingent. For most of the survey items, they didn't differentiate much from the rest of the sample but they did for a couple of important ones. Fully 74% of these instructors reported having received SETs comments that they perceived as hostile, abusive or bullying in comparison to 58% of tenured/tenure-track and of those, 41% reported feeling unsafe as a result of receiving such comments in comparison to 24% of tenured/tenure-track.

[Slide 35]

Wrap-up and key findings

Our study also shows that there is a considerable proportion, almost 2/3, who receive hostile and bullying comments. This is exacerbated in several major equity seeking groups.

There are considerable negative wellbeing impacts which again correspond with ascribed traits of the instructors.

Concerning the pedagogical impact, the findings show a marked incongruence between what SETs comments are **ostensibly** designed to do and what they actually do.

[Slide 36]

There are studies which try to nail down the exact frequency of hostile and bullying comments in any batch of SETs comments, and then decide whether they attain some threshold level we should be concerned about. They tend conclude that in fact there's no problem with SETs comments. Apart from methodological problems in those kinds of studies, we consider that to be a pointless exercise.

I say that partly because our open-ended responses show that **hostile** comments are more impactful and memorable than positive ones.

This remark on one of the open-ended items is illustrative:

*"The term that I had some students complain about my asthma-related coughing, and some others complain about my apologies for coughing, was the term that my attention to SETs began to wane significantly. I find there is little of value--and much that has the potential to hurt and disappoint--that comes out of giving hundreds of people an anonymous platform to say anything they want about me with impunity. I feel this way despite the fact that, overall, I largely get positive reviews. **The negative comments are the ones that stick with you.**" – Phil, 44*

[Slide 37]

Recommendations

We concur with several other recent critiques of SETs overall, including OCUFA in its 2019 report, that SETs comments should only be for formative not summative purposes.

We think that there needs to be more attention to how the comments, not just the scores, are processed, in semester-end evaluations. The comments matter.

Having served on tenure and promotion committees numerous times over my career I can tell you that people do notice and comment on negative student feedback, in assessing an instructor's performance; they can influence merit ranking, renewal, and so on. Yet there is so much that can't be known about the context of those comments.

The second and third recommendations are corollaries of the first. Formative purposing of student feedback requires that instructors engage in an ongoing teaching and learning relationship with their students throughout the course. Faculty can use those to demonstrate to their colleagues evaluating them that they have done this, and what insights were gleaned from it.

[Slide 38]

Students should be given guidance and instruction about how their comments are used. OCUFA argues there's a research ethics aspect here to be considered, when we consider SETs scores and comments as a form of data being collected: if it's not fully explained to students **how** the data will be used, then they cannot give informed consent.

As well, there is data to show that when it **is** carefully explained to students why and how their feedback is utilized, how their biases can affect it, how the recipients of inappropriate and hurtful comments are affected by that. The feedback students give is more respectful and more useful, than without that kind of intervention.

Regarding supplementary forms of evaluation here especially we have in mind peer observation. I commend my own University for offering that option, starting this year, through a Letter of Understanding with the faculty association.

Suspension of SETs for contingent faculty

We would love to hear from contingent faculty on this point. We think that they would benefit as much as tenure-track faculty from a transition to formative purposing of SETs comments and furthermore, there are too many ways that poor SETs can hurt contingent faculty, whereas they do not stand to benefit in anything close to the same way as their tenured or tenure-track peers from excellent ratings.

[Slide 39]

Times are a changing

We should acknowledge a couple of recent events that may **already** be driving changes in how universities do SETs which may be opening opportunities for change.

One is the June 2018 Ryerson arbitration that I mentioned earlier. Many people predicted ripple effects from that, across North America, and there have indeed been changes in some universities; I mentioned my own university in this regard.

However, we know of no overview study of how much this kind of change is happening. I think we need such a study, quite urgently. From having spoken to a few colleagues at various different universities last year, I suspect that willingness to change SETs is uneven.

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

Secondly, COVID-19. The pandemic is likely to change the weight assigned to SETs and even whether they are implemented at all, certainly in the short term. It's not clear how that will settle out in the medium and long term. At least, to me it's not clear, but some participating in the webinar might have informed predictions to share at the end.

These shifts create an opportunity for fulsome change in how open-ended feedback from student is sought and how it's used. I want to reiterate again that while there's been considerable attention to changing how we handle numerical ratings of teaching are obtained and processed, we must not neglect the qualitative comments. I look forward to the discussion.

Finally, I want to thank the CSA for providing most of the funding the research and we also thank OCUFA their generous contribution.

I especially want to thank Sherry Fox, the CSA's Executive Director, for her invaluable support to us throughout the project and a huge thanks to you Rochelle for moderating, and for your very helpful input into the study design some months back.

[\[Questions and Answers\]](#)

Rochelle:

Thank you, Lisa. I can't say enough good things about the report and know how much time and energy you and your colleagues (Rochelle and Rachel) put into this study and the production of this report and you should all be very proud of yourselves.

I'll now go into the next part of this report which is questions from today's attendees to see what they have to say and what they are curious about. Just by way of housekeeping, I would like to remind everyone that there is the Q & A function at the bottom of your screen that you can use to ask questions.

Also, in case you hadn't noticed, if you look in the top left-hand corner of your screen, you'll notice that we are recording this webinar so if you have to leave early or if you are looking to revisit this talk or the questions from today, you will have the opportunity to do so. The CSA will be posting this on their website.

I had some questions prepared but I think I will just go into the Q & A to have a look at what our audience is wanting to ask and I will save my questions for Lisa for a little bit later.

For the faculty who got lower rating due to her age (who no longer showed care about her students), might this also be an indicator that she might want to re-energize her teach practice. Introducing new methodology? I've worked with faculty who refuse to improve their practice over the years.

Lisa:

That is a really interesting comment. I am sure it could be the case that faculty (both young and old) could be less energetic than we would want them to be about their teaching. Maybe I am naïve but I tend to assume that people across all age ranges that I know in my department and elsewhere take great pride in their teaching and they do all they can within their time availability to energize. I think it is always a good idea for any of us to consult with experts on teaching to overhaul our design and so on and also be able to narrate that we say to people evaluating us in our teaching dossiers. We want to demonstrate 'These are the things I did. I consulted and went to redesign institute.' I tend not to believe that older faculty are less energized/less energetic and I say that as I am getting there myself. I feel that I am constantly learning about teaching and open to improve it. I can't see a difference in colleagues who are older than myself.

When I was a student at both McMaster and York University, we had a focus groups with a T&P committee member after they had completed a teaching observation as part of the tenure process for faculty members. In both instances the focus-group session was transcribed and combined with the observations to create a report-like document. Is there any evidence that alternative qualitative practices like this might be superior? Or, how widespread these alternative practices might be?

Lisa:

I think that we need to see more research on what happens with alternative methods of teaching evaluations. How does it affect teaching when we have peer observation? How does it affect teaching when we give students lots of good advice about the meaning of what they are doing? I don't think we have enough research on that and we do really need that kind of research. I can't say for sure that there is evidence that it is superior or not. I think that we need alternatives to what is being done right now and I think that instructors who do that kind of process you are describing should be proud of that and marry that with they report about their teaching.

I am very encouraged by this study. I am wondering what might be suggested for contract faculty to replace the SETs within the Teaching Dossier?

Lisa:

My sense is that it would stand to benefit by making SET comments and scores formative rather than summative. I think that there is too much evidence that contingent faculty can be hurt by negative scores and even a few outlier negative comments. How exactly this would work and what it would look like is a question that is very important. Our set of recommendations is a step towards fully working this out but I would like to see contingent faculty working with tenured faculty to devise a system that is better for them. It needs to have their voices and experiences. There are people working on this topic. I mentioned the research study by Smele et al. which I think is opening the door in that direction. We can then devise something that is much more useful and would be evidence on ways people are approaching their teaching without bringing summative scores and comments.

I am wondering what your thoughts are regarding the idea that qualitative comments offer some context to the quantitative numbers on course evals. In this sense, if an inappropriate comment were to be made and this corresponded to a low eval score we would have good justification to drop the entire course eval from the instructors' ratings. Without the qualitative comments it can be hard to make sense of the numbers. I feel we are missing some context that could help.

Lisa:

That may be the case but I also want to put out there the possibility that comments instead of providing context may actually be missing context. Comments faculty receive may not be showing what was happening in that course. Apart from the size of the course and so on, could it be that the topic the course was teaching on was considered too controversial or complicated for the students. There are some cases in which the faculty member may be dealing with situations of contra-power bullying or harassment in the classroom or social media settings. So, I think that sometimes the comments are missing those contexts. So you may be right but also comments themselves need to be contextualized.

Could you comment more on themes emerging with LGBTQ instructors/faculty?

Lisa:

I wish I could but our survey just slightly delved into how there is correlation between the items we are interested in and whether someone is heterosexual or LGBTQ. The literature on this is also contradictory. Some of the literature we cite in our report shows that some LGBTQ faculty are getting more negative scores or there is no difference.

I should also say that in our open-ended feedback, we also found a mix of things. Some faculty were seemingly being condemned for being openly gay and others were considered heroic for being openly gay in their teaching. I feel like the literature and also our open-ended comments are showing contradictions on this. Further, fine-tuned, research is needed on this to really understand how this is playing out.

Excellent report Lisa - kudos to you and your colleagues. I'm wondering - given what you have learned from faculty members, do you have recommendations for instructors around the kinds of evaluations they might implement themselves? Perhaps one at the beginning of the term, then others throughout? (This is similar to the question just asked.)

Lisa:

I think there are all kinds of interesting possibilities with soliciting formative feedback. I think that comments matter more than scores. It is possible to use a rating system throughout the semester. At the beginning and the end and also at different points throughout the semester depending on what is the aspect of the course design or the assignment that the instructor wants to find out about. For our own formative development,

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

it is a really good idea to do that. Also for how we demonstrate that to our peers evaluating us to show that we did that and indicate the insights that came out of it.

I think that how the questions are worded depends very much on the specific course. Giving faculty members control over that (the writing of the questions) is extremely important as well as the timing on when to do that.

Rochelle:

We have another question from an anonymous attendee. Thank you for taking the time to ask your question even though you may not feel comfortable with giving your name. I would like to encourage other attendees for the webinar to feel free to ask your questions anonymously if you are not comfortable giving your name.

Can you speak more about how racism can affect the way a student approaches the SETs (i.e. accents, different teaching styles based on international training, BIPOC courser content), and how this is considered (or not) by tenure and promotion committees.

Lisa:

The literature is pretty unequivocal about how racism among students - reflecting racism in the broader society - plays out in SET scores. There is study after study, especially over the last decade or so that make it practically incontestable that there is a problem in how racialized groups are being evaluated and how their scores are lower. What about the T & P committees?

Perhaps I am being naïve in suspecting that a certain minimum will be cleared during the process which I can certainly say within my own university. We are required to do anti-bias training before we take part in the T & P process. This requires us to think about exactly all of these things.

I would hope that minimum bar is cleared throughout across various institutions. Do I know this for a fact - I don't. It would be good to have research that looks into this.

Rochelle:

As far as I know, my university does not require us to take any diversity or cultural sensitivity training before being on a T & P committee and I commend your university for doing that.

Thank you very much for the work. I find it unfortunate that this practice has been ongoing since the 1970s and not much study has been done yet the impact is real and especially negative for equity seeking groups. Are there resources websites where alternative forms of feedback exist?

Lisa:

I don't know if there are websites that exist however I have been impressed by research from an American expert on post-secondary education, named Cook-Sather, which looks into

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

incorporating students as peer reviewers but in a different way than conventionally done. Students were not taking the actual class but were trained and paid to do objective observation of the course. I think her work is a great resource to consult about this. I am not aware of a website but maybe others participating may know. If it doesn't exist, maybe it is time to put one together which could be particularly useful for the Canadian context.

Rochelle:

Maybe that is something that we can encourage the participants to do. If they do know of alternative websites that provide other methods of observation, that they post them in the chat box or in the Q and A. We will make sure that they are published as part of the webinar transcript.

I'm curious about what guidance can be provided to P&T committees and Deans re SETs. Some seem to think that those who get below average scores are poor teachers, but an average is an average - it is not a standard. Are there any universities that set targets/standards rather than pit faculty against each other through the use of averages? And that's just about the numeric scores - the broader issue is what are better ways to assess quality (like peer reviews, etc.) which you have begun to address.

Lisa:

Two things I want to say about this. One is that we actually need a study to see what is going on at universities post-Ryerson ruling. What kinds of changes are happening with the treatment of scores and comments?

I have to give kudos to my own university. Their instruction is to not compare averages to scores across faculty. The only scores that are relevant is comparing the professor to themselves. Not to look at averages - especially considering the absurdity of some of the questions like 'How stimulating was the person?' I suspect that we are not the only university doing that but it would be important to look at what is being done across the country and North America.

Rochelle:

Thank you, Lisa. So, I am going to take my chair privilege to ask a question. Throughout this webinar and question period, people may be wondering what are the takeaways for faculty, lecturers and those actually teaching. We have another question about what university administration T & P committees can do with the findings of this report. One of the things we may not think about is the potential benefit of this study to students going forward. Can you speak to how you think this report could be helpful for students?

Lisa:

Students can only benefit from this long-overdue overhaul how we process the SET scores and comments. I think that doing so is not going to put an end to negative feedback. In fact, we are likely to get negative but constructive feedback on courses that can only be helpful to students. I also think that for too long, university administrations have neglected a duty they have to advise students about what all this means. We are then faced with a social media

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

trolling culture outside academe but the boundaries are permeable and finds its way into the private sites (like Rate My Professor) and weaves into the official evaluations as well. Students are going to find themselves in the same situation someday - rated like an Amazon product. For them to obtain the skills in doing this, it will help them to give meaningful and constructive feedback. It will rebound in better teaching because it will liberate faculty to dedicate their energies without fear of falling on their faces if a new innovation does not go over well. We see in the study that some people are afraid of innovation because they are worried about negative feedback. We can liberate those energies and seek student feedback throughout the course with a real relationship with students without leaving it all to the end. This avoids the 'halo effect' whereby one over-arching question affecting all the smaller questions. Similar things may happen in the open-ended comments. I believe it is a win-win situation for both students and faculty.

What types of strategies would you suggest to help train students how to fill-out course evals in more meaningful ways? I know McGill has some resources but just curious if you have some ideas (We are working at UW to help inform students and have found in our own research many lack an awareness that course evals are used summatively)

Lisa:

I will refer to a report we cited which looked at nursing faculty (Moralejo et al 2019). Although they had only five faculty, they had hundreds of students in the study. They ran an experiment in which they had one group which was provided training and guidance on what all of this means and one group which did not receive that intervention. They found an important difference between the groups whereby the ones that received intervention actually gave more negative comments but did it respectfully and constructively. It would be good to refer to that study and talk to them about it to see how they did that. It could serve as a model or starting point to build that up.

Rochelle:

We are going to take just one more question unless there is anybody else out there who has other burning questions to ask of Lisa. We will keep the Q & A open for a few more minutes and if we do not have time to answer everyone's questions, Lisa's responses will be included in the transcript of today's webinar.

Your study focussed on Sociology profs, and it's been my experience that Humanities and SS instructors are expected to be more entertaining than instructors in the "harder" sciences. It would be interesting to see how this plays out over Faculties and schools.

Lisa:

I don't know about that but there are a group of scholars at my own university, led by Omer, who have a forthcoming study which focuses on open-ended feedback comments. They compare STEM profs to social sciences and humanities profs. It doesn't quite address what you are getting at but they did find that STEM faculty were overall more satisfied with SETS comments than non-STEM faculty. It would be worth looking into. Maybe their study will look at whether the 'entertainment' aspect has any significance. I suspect that there is an

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

expectation that even in the STEM faculty are positive, energetic, passionate, happy. I will note the OCUFA report from 2019. There is a study by a University of Waterloo Professor which looked into what is reflected in high SET scores. Much of it is related to personality but may not distinguish between STEM and non-STEM faculty. You may want to look into that report.

Rochelle:

Do you see your study as the final word on qualitative comments and SETs or more of a starting point?

Lisa:

Speaking on behalf of my co-researchers as well, we do not see this as a final word. There are six recommendations that is interwoven with what was found in other literature and we hope for engagement with people who are studying this, people who are experts in education in sociology and beyond. It would be terrific if they could give us some viewpoints on our recommendations - on the content of them and how to make them impactful.

Through this engagement, we hope to produce a compound document and set of recommendations based on ours as well as other experts working on this issue.

Rochelle:

We will do one last question for today. Thank you so much everyone for your participation and sharing questions with Lisa. Thank you again Lisa for giving us this excellent summary of the report.

One additional comment below;

I think the negative effects of SETs cannot be entirely undone, given the availability of online media to students to do the same. And, they have positive effects as well. So, perhaps we should think about supplementing them with peer evaluation/feedback + giving instructors more control over the questions/contents + instructor's decision about what they should be used for and which parts...

Lisa:

Supplementing is a great step forward from where we are now although there has always been some in the teaching dossiers. Nevertheless, I think summative scores and comments have an unwarranted impact and influence on your colleagues' evaluation of your teaching.

The reason we recommended replacing (and we concur with OCUFA on this) the summative scores and comments is because they still have that negative impact. Even with training and guidance for the T & P committees, there is always that hazard. With formative feedback, it won't rule out racially biased comments but it would not have the same consequence for contract renewal, tenure, promotion, and so on. Social media is always out there where people can say whatever they want with impunity. I don't know how to deal with that

(Re)Prioritizing Pedagogic Feedback:

Faculty Experiences with Qualitative Comments from Student Evaluations of Teaching (SETs)

particular issue but I feel that the administration can step up to its duty of making students understand how important this is and to be constructive and respectful.

Rochelle:

I would like to draw everyone's attention to the screen where it gives you the web address to be able to download the report that Lisa, Rachel and Rochelle worked so hard on. If you have any questions or comments, contact Sherry at the CSA or office. Check back on this website for a copy of the webinar recording and transcript.

Thank you everybody. It has been a true pleasure, Lisa, having a chat with you and also looking at all of the questions and comments people had today.

Lisa:

Thank you everyone for your comments and participation.