

Winter 12-31-2019

## Measuring Student Responses in and Instructors' Perceptions of Student Evaluation of Teaching (SETs), Pre and Post Intervention

Lisa Moralejo

*University of British Columbia Okanagan, lisa.moralejo@ubc.ca*

Elizabeth Andersen

*Thompson Rivers University, eandersen@tru.ca*

Norma Hilsmann

*University of British Columbia Okanagan, norma.hilsmann@ubc.ca*

Lindsay Kennedy

*University of British Columbia Okanagan, lindsay.a.kennedy@ubc.ca*

Follow this and additional works at: <https://www.cjsotl-rcacea.ca>  
<https://doi.org/10.5206/cjsotl-rcacea.2019.3.8052>

---

### Recommended Citation

Moralejo, L., Andersen, E., Hilsmann, N., & Kennedy, L. (2019). Measuring student responses in and instructors' perceptions of student evaluation of teaching (SETs), pre and post intervention. *The Canadian Journal for the Scholarship of Teaching and Learning*, 10(3). <https://doi.org/10.5206/cjsotl-rcacea.2019.3.8052>

---

# Measuring Student Responses in and Instructors' Perceptions of Student Evaluation of Teaching (SETs), Pre and Post Intervention

## Abstract

At most colleges and universities, students are invited to complete Student Evaluation of Teaching (SETs), which have both formative and summative purposes. In this convergent mixed methods study we evaluated if we could influence (a) students' numerical responses and nature of their comments and (b) instructors' physical and emotional responses to SET results, their perceptions of their results, and perceptions of SETs overall. Students who received an in-class intervention submitted more qualified comments, addressed specific issues, and made more recommendations for improvements compared to students who did not receive the intervention. Instructors reported reduced physical symptoms related to SETs after they received the intervention. Instructors reported that the intervention helped them let go of feelings of frustration and isolation and that they had acquired new strategies for opening, reading, and interpreting SET results. They continued, however, to report feeling apprehensive, uneasy, and uncertain about impending SET results.

Dans la plupart des collèges et des universités, les étudiants sont invités à remplir des questionnaires pour évaluer l'enseignement, dont les objectifs sont à la fois formatifs et sommatifs. Dans cette étude axée sur une convergence de méthodes diverses, nous avons évalué si nous pourrions influencer (a) les réponses numériques des étudiants et la nature de leurs commentaires, et (b) les réponses physiques et émotionnelles des instructeurs aux résultats de ces évaluations de l'enseignement par les étudiants, leurs perceptions de leurs résultats et leurs perceptions de ces évaluations en général. Les étudiants qui avaient obtenu une intervention en classe avaient remis des commentaires plus qualifiés, avaient traité de problèmes spécifiques et avaient proposé davantage de recommandations pour des améliorations, en comparaison des étudiants qui n'avaient pas obtenu d'intervention. Les instructeurs ont rapporté qu'ils avaient moins de symptômes physiques liés aux évaluations de l'enseignement par les étudiants après avoir obtenu l'intervention. Les instructeurs ont déclaré que l'intervention les avait aidés à se débarrasser des sentiments de frustration et d'isolement et qu'ils avaient acquis de nouvelles stratégies pour ouvrir, lire et interpréter les résultats des évaluations. Toutefois, ils continuaient à avoir des sentiments d'appréhension, à se sentir mal à l'aise et à connaître une certaine incertitude à l'approche de recevoir les résultats des évaluations de leur enseignement par les étudiants.

## Keywords

teacher evaluation, student feedback, university, instructor; évaluation des enseignants, feedback des étudiants, université, instructeur

## Background

At most colleges and universities, students are invited to complete Student Evaluation of Teaching (SET) every term. Universities and colleges typically use SETs as a measure of teaching quality (Chan, 2001; Chen & Hoshower, 2003; Roberts, 2015; Spooren, Brockx, & Mortelmans, 2013). Therefore, the evaluations have both formative (to promote professional development) and summative (to hold teachers accountable for their practices) objectives (Chan, 2001; Delvaux et al., 2013).

While most teachers support SETs for formative purposes, many are concerned about the summative purposes and feel anxious and apprehensive if decisions are made regarding tenure, promotion, and salary increases based on SET results (Beran & Rokosh, 2009). In a precedent-setting case, an arbitrator ordered that SET results not be used to measure teaching effectiveness for tenure or promotion at Ryerson University in Toronto, Canada (Farr, 2018). Additional concerns with SETs include: doubts about reliability and validity of SET instruments, reservations about students' abilities to provide appropriate evaluation, depersonalization or apparent lack of student awareness regarding the importance of their online responses, students' lack of motivation to complete online evaluations, and insufficient evidence to support a positive correlation between student ratings and actual student learning (Beleche, Fairris & Marks, 2012; Chen & Hoshower, 2003; Spooren, Brockx, & Mortelmans, 2013; Uttl, White, & Wong Gonzales, 2016).

Concerns and doubts about the validity of SETs can directly impact utility of the feedback. For example, if student responses are limited (small sample size), some teachers may interpret the responses as unrepresentative of the entire class or representative of students with extreme opinions only (Winer, DiGenova, Costopoulos, & Cardoso, 2016). Some teachers may perceive that certain courses in the same discipline and/or certain disciplines routinely receive lower SET scores compared to other courses or disciplines. There may be some evidence to suggest that their perceptions may be true (Gravestock & Gregor-Greenleaf, 2008). While others may perceive that larger class sizes, timing of administration of SETs, and "rigorous" grading practices all negatively impact the student responses and result in lower SET scores (Winer et al., 2016). Finally, researchers note that some teachers focus most on student complaints and subsequently experience feelings of depression, devaluation, and dejection (Moore & Kuol, 2005). Although Winer et al. (2016) suggest that the evidence does not support the majority of teacher concerns, Darwin (2016) submits that tension surrounding SETs is growing and as a result, negative teacher reactions to performance feedback may have a direct bearing on their subsequent efforts to improve, sustain, or enhance their teaching performance in future courses. Few researchers have examined if students' callous comments in SETs can be reduced and their solicitous behaviors enhanced, or if instructors' perceptions of SETs and/or the nature of their reactions to SETs can be transformed (Yao & Grady, 2005).

This convergent mixed methods study focused on the perceptions and experiences of university instructors who received SET results from undergraduate students enrolled in their courses. There were two overarching aims. First, to evaluate if an intervention for undergraduate students would influence their numerical responses on the SETs for a particular course and the nature (tone and content) of evaluative comments. Second, to evaluate if an intervention for a group of instructors would influence their perceptions of SETs overall, their perceptions of their own SET results, and their physical and emotional responses to the results.

## Method

### The Context

This study took place at one university in the province of British Columbia, Canada, within a School of Nursing where approximately 54 instructors teach over 500 students enrolled in the four-year baccalaureate nursing program. Approximately 50% of the students are admitted directly from high school and 50% are university transfer students who have completed 24 credits of university level. Approximately 10 to 15% of the students are male, and 85 to 90% are female. While the scope of this mixed methods study does not allow for generalizability of the results, the results may be transferable.

### Recruitment and Sample

After we received administrative support and approval from the Behavioural Research Ethics Board we employed purposive sampling to recruit one experienced instructor who represented a unique and “information rich case” (Sunderji & Waddell, 2018, p. 1093). Then we employed convenience sampling (Strauss & Corbin, 1998) to recruit additional participants.

To be eligible as an “information rich case,” an experienced instructor had to be teaching an undergraduate course divided into two equal sections (Section 001 and Section 002), willing to share numerical SET results and student comments with the researchers, and willing to participate in two semi-structured interviews and three group discussions. To recruit the “information rich case,” names of eligible instructors were obtained from an Educational Plan distributed to all teachers in the School of Nursing. Instructor email addresses are publicly available on the School of Nursing website. We emailed all eligible instructors and explained the general nature and purpose of the study and the time commitment required. We attached an information sheet and consent form, but it was up to individual instructors to contact us.

Convenience sampling was used to recruit additional instructors willing to participate in two semi-structured interviews and three group discussions but not required to share their numerical SET results and student comments with the researchers. To be eligible to participate in the group discussions, instructors were required to be teaching in the School of Nursing and able to reflect on their experiences receiving SET results from at least one theory or clinical practice course. Potential participants for the discussion groups received our information letter and invitation to participate via a school listserv and were instructed to contact us if they were interested in participating and met the eligibility criteria. Four additional instructors agreed to participate. The characteristics of all instructors ( $n=5$ ) are listed in Table 1.

Table 1  
*Instructors' Characteristics*

Age	Gender	University Teaching Experience	Level of Education
> 50 years = 2	Female = 4	> 15 years = 1	Master's degree = 4
40-44 years = 1	Male = 1	10-15 years = 2	Bachelor's degree = 1
30-39 years = 2		2-5 years = 1	
		7-10 years = 1	

We sought and received ethical permission to proceed without student consent. The research involved no more than minimal risks to the students and the alteration to consent requirements was unlikely to adversely affect their welfare. Given the research design, we anticipated that student SET comments and numerical responses would become skewed if the students knew about the study ahead of time.

The final sample, therefore, consisted of one experienced instructor who was the “information rich case” ( $n=1$ ), additional nursing instructors ( $n=4$ ) and undergraduate nursing students ( $n=128$ ) enrolled in two sections of a first-year professional development course (Section 001 or Section 002).

Qualitative and quantitative data were collected concurrently. All instructors ( $n=5$ ) met individually and in private with the principal investigator to describe their overall perceptions of SETs and their experiences receiving SET results. These interviews were semi-structured, lasted approximately 30 minutes, and were audio-recorded to enable qualitative analysis. Instructors completed a 28-item Likert scale questionnaire developed by the researchers and based on physical and emotional responses to SETs that were known to researchers or highlighted in the literature about similar teacher evaluations. Finally, they were invited to attend three lunch-hour group discussions facilitated by the researchers. Group discussions were focused on providing peer support and time for contemplation of experiences receiving SET feedback. Group discussions were not audio-recorded, but the researchers took notes for later qualitative analysis. All interview and discussion group data were collected in the instructors’ natural work environment.

Students enrolled in section 001 of the “information rich case” instructor’s professional development course ( $n = 59$ ) received standard course content. Students enrolled in section 002 ( $n = 69$ ) received an intervention in addition to standard course content: three 15-minute presentations. The presentations were designed to enhance their abilities to provide effective feedback to peers and others and were scheduled approximately 3 weeks apart (Table 2).

Table 2  
*Student Presentations*

	Key Objectives: At the end of the presentations, students will:	Characteristics of Presentations
Presentation One	Describe the development of SETs Explain the value of SET results Discuss the use and potential impact of SET results for instructors (tenure, promotion, salary)	PowerPoint
Presentation Two	Contrast constructive and non-constructive feedback Describe negative personal repercussions for instructors who receive offensive, inappropriate and/or unconstructive SET comments Consider how online anonymity can negatively affect comments	Visualization exercise: students asked to imagine receiving negative feedback from an instructor at the end of a clinical practicum without any prior feedback  Interactive discussion after actual examples of non-constructive/acerbic SET comments were read aloud
Presentation Three	Describe approaches to providing constructive feedback Apply the principles of providing constructive SET feedback	PowerPoint: Tips on providing effective feedback Video of a teacher using ineffective teaching practices and asking students to provide writing constructive feedback for this teacher

Students enrolled in both sections received identical instructions to complete SETs and were provided 15 minutes in class to complete the SETs online. Following usual SET practices, students also responded in unlimited text form to three questions: What were the strengths of this course? What did you most enjoy about it? What were the weaknesses? During the analysis, student responses to the questions “What were the strengths of this course?” and “What did you most enjoy about it?” were categorized as positive comments and student responses to the question “What were the weaknesses?” were categorized as negative comments.

The Centre for Teaching and Learning at the university manages the SET data collection for the entire university. That is, they distribute the SETs to students online, tally the numeric results, collect the student comments, and distribute the results to instructors. The SET results are anonymous. The SET data that we accessed, therefore, was secondary data. The “information rich case” instructor who had taught the first-year undergraduate professional development course shared both sets of numerical SET results and both sets of student comments (section 001 and section 002) with the researchers.

The principal investigator met with all instructor participants again (individually) to discuss their experiences and their physical and emotional responses to reviewing their most recent SET results. These interviews were also audio-recorded to enable qualitative analysis and the instructors completed the same 28-item Likert scale questionnaire focused on physical and emotional responses. One instructor did not complete the post discussion group interview.

## Confidentiality of Participants

Confidentiality of instructors who participated in the discussion groups was maintained through the use of pseudonyms and all participants were asked to keep what was discussed during discussion groups confidential. Although researchers are bound by ethical responsibilities to maintain confidentiality for participants, full confidentiality could not be guaranteed on behalf of the other discussion group participants. This limitation was clearly noted on the consent form. During dissemination of the results, the students and members of the School of Nursing knew the identity of the “information rich case” instructor who taught the professional development theory course, so confidentiality of that particular instructor was not maintained during local dissemination. The SET results were anonymous, so individual students were never identified in this study.

## Data Analysis

Using a convergent design the data were analyzed in three separate and distinct stages (Creswell & Plano Clark, 2018). First, the qualitative data were analyzed independently. Notes from the three discussion groups were analyzed using a general inductive approach (Thomas, 2006). Pre-discussion group themes were compared to post-discussion group themes. The nine semi-structured interviews with instructors (five pre-discussion groups, four post-discussion groups) were transcribed verbatim and analyzed using the same general inductive approach. To maintain rigor, we concentrated on what the participants said, being open-minded and responsive to their statements, and avoiding any preconceived assumptions. All four researchers met to share work and verify that the progression of codes and themes were valid.

Next, the quantitative data were analyzed independently. The SET mean numeric responses from students in section 001 were compared to the mean numeric responses from students in section 002; the nature (tone and content) of the students’ evaluative comments for sections 001 and 002 were compared using a rubric developed by the researchers. Finally, pre-and post-discussion group data from the 28-item physical and emotional response questionnaire for instructors were compared. Due to the small sample sizes, the statistical results are descriptive rather than inferential. As a last step, the quantitative and qualitative data sets were merged to identify content areas represented in both data sets (Creswell & Plano Clark, 2018).

## Results

### Qualitative Results (Repeated Group Discussions with Instructors)

Four themes emerged after written notes obtained during the three discussion groups with instructors were analyzed qualitatively: The Foreboding Shadow; That’s Going to Leave a Mark; Coming to Grips with Lack of Control; and Creating a Space and Letting Go.

The Foreboding Shadow reflects instructors’ apprehensive feelings towards receiving SET results. All of the instructors shared anxieties and enduring beliefs that SETs are likely to reveal at least some alarming or frightful results. Most of the instructors described how they delayed reading the SETs until they felt ready to confront unpleasant comments. Some delayed reading SET results for months. Others waited until they felt ready and had found a private space. Some

described bolstering themselves with alcoholic beverages before reading: *It's like when you get the Visa bill after Christmas.*

Feelings of dread set the stage for the second theme 'That's Going to Leave a Mark'. Instructors agreed that some student comments made it difficult for them to continue to feel enthused about working and teaching: *I had a very hard time getting to the classroom after reading the SETs.*

Regardless of how many positive comments they received, instructors shared a common tendency to focus most on the negative comments thereby allowing a few negative comments to overpower or override the benefits of multiple positive comments: *It takes so many positive comments to counteract the negative ones.*

Ultimately, instructors described feeling distressed, overwhelmed and tormented by negative comments: *It blows my mind that students have an avenue to bully instructors.*

All participants expressed feeling frustrated with one particular aspect of SETs - inability to address student complaints, respond to, or counter negative student comments. Becoming accustomed to long term feelings of helplessness and defenselessness as a result of inability to counter is illustrated in the third theme 'Coming to Grips with Lack of Control'. Some instructors described feeling especially helpless if negative student comments focused on classroom environmental challenges such as heating, lighting, seating or class scheduling options: *Yes, the room was too cold, too hot, too crowded, too far... and I don't want to be here on a Monday evening or a Friday afternoon either!*

Other instructors described feeling constrained by teaching assignments that included an unpopular core course: *It wouldn't matter if Beyoncé came and taught the research course!!*

Instructors also wished that they had the ability to respond to students whose negative comments suggested that they misunderstood or misinterpreted part of the course content: *I believe that I may have mentioned that a zillion times?*

Ultimately, instructors agreed that intermittent feelings of helplessness and defenselessness are a normal part of the SET process. Metaphors for SETs that were used during discussions included: a festering wound, a fish hook, a ticking clock, a ticking time bomb, an attack, and a Visa bill.

'Creating a Space and Letting Go' reveals that attitudes of defenselessness and helplessness towards SETs can be shaped by socialization and peer support. Most instructors suggested that they felt less isolated and alone because group discussions offered a place to express perceived injustices. They indicated that they enjoyed their peers' revelations and felt relieved and enlightened to learn that their peers had similar responses, coping mechanisms and feelings. Some described making conscious decisions to change what felt detrimental and debilitating (such as feelings of apprehensiveness and lack of control) by making the most of constructive feedback, and letting feelings of depression and frustration go: *Instead of just festering - take the fishhook out and let it go.*

### **Qualitative Results (Semi-Structured Interviews with Instructors)**

After the semi-structured interviews were analyzed qualitatively, ten themes were revealed. Pre-discussion group themes from semi structured interviews with instructors are compared to post-discussion group themes in Table 3. Most of the themes generated from pre and post interviews were similar to the themes generated from analysis of the discussion groups. For example, the theme 'Why do today what I can put off...until much, much later?' mirrors the



procrastination uncovered in discussion group theme ‘The Foreboding Shadow’. Three of the ten themes, however, were different: ‘Houston, We Have a Problem’, ‘I act, I Perform for SETs’ and ‘The Dirty Little Secret’.

‘Houston, We Have a Problem’ is a reflection of some instructors’ beliefs that a significant and unaddressed problem occurs because SET responses are anonymous. These instructors believe that the anonymous feature of SETs enables a small minority of students to use them for vengeful purposes or to right perceived wrongs. These instructors described SETs as opportunities for some students to cyber-bully and suggested that universities should review student comments and filter them before releasing them to instructors: *I know at the end of the term students can have their payback.*

While most instructors use SET results to improve their teaching practices, the theme ‘I act, I Perform for SETs’ reveals that instructors sometimes find themselves catering to students in order to improve their SETs. The SETs therefore have potential power as pedagogical constraints. Some instructors described intentionally sacrificing some elements of their personal integrity, reluctantly succumbing to unreasonable student demands, or manipulating their instructional presentational norms and behaviors because of impending SETs: *I almost cater to the SETs.*

‘The Dirty Little Secret’ draws attention to self-stigma. During the individual interviews, instructors revealed that they did not feel adequately prepared to receive, interpret, and utilize SET results. Some instructors reported suffering in silence for years, accepting negative and often inaccurate student comments about their teaching abilities. These instructors revealed that they had internalized negative student comments and had never talked to other instructors about the SET process: *None of us were prepared to get unsolicited and unprofessional and uninformed feedback.*

For the most part, analysis of the semi-structured interviews with instructors revealed that participation in supportive peer discussions could positively influence subsequent willingness to identify and change internal beliefs, values, and opinions about SETs. Most participants revealed that they were able to reduce the stress, trepidation, and procrastination associated with SETs and replace old strategies for opening, reading, and interpreting SET results with new and more effective systematic strategies. Participants confirmed that no instructor is perfect, no one can please all students all of the time, instructors need not feel alone in self-stigma, some student comments are neither constructive nor useful, and the ability to distinguish between constructive and unconstructive comments will help instructors to focus most on the meaningful or constructive feedback.

### **Quantitative Results (Students)**

- Mean numerical student responses to each survey item on the SETs are provided in Table 4. The nature (tone and content) of the students’ evaluative comments are summarized in Table 5.

Table 3  
*Pre-Discussion Group Themes compared to Post-Discussion Group Themes*

Semi structured Interview Themes Pre-Discussion Group		Semi structured Interview Themes Post-Discussion Group	
Theme	Illustrative Quotes	Theme	Illustrative Quotes
	Do I really need to look at them or could I just maybe pretend that they don't exist and live in the land of denial for a while?		For the first time ever I opened them right away. I read every comment, then closed it and then went back to it again. I felt very much at ease
			Warm cup of favorite tea in hand
Why do today what I can put off until... much, much later?	I would never actually read the comments during the business day because they can be so hurtful and it would actually impair my ability to do a good job	To cope or not to cope isn't the question...but how?	Having participated in the discussions helped me to sift through some of the emotional part and unpack it a little bit
			I actually learned (from listening to all the other instructors) to not just open it up and sort of willy-nilly go through everything. Instead, I systematically looked at the overall graphs and then looked at the positives and then, when I felt like I had a good feel of what was going on, looked at the "needs improvement" area
			I tried to focus more on what was relevant, recognizing that some comments really aren't grounded in anything constructive
			I didn't drink this time

Internalization	<p>It feels a little like getting a package in the mail that is ticking</p> <p>I get extremely anxious before and experience physical symptoms of stress</p> <p>I always struggle with opening the SETs in the middle of a working day and then have to continue to be an effective teacher</p>	I am not alone	<p>I didn't realize that there are others having very extreme reactions. Not just one or two but quite a few</p> <p>I think that I felt less apprehensive after all of our group sessions; like I somehow wasn't "alone"</p> <p>I didn't take it to heart as much. I tried to make it a positive</p>
Houston, we have a problem	<p>I know at the end of the term students can have their payback</p> <p>Students' ability to respond anonymously is a little bit like cyber bullying</p> <p>Not a lot of thoughtfulness that there's a real person on the other end reading what you're saying, just a place for students to vent</p> <p>Students are out to right something where they have been wronged.</p> <p>I find that students who are mad or happy respond but those who are just content don't</p> <p>Some courses are more popular or students see them as more important, so there are some courses that get more negative SET results.</p>	Teaching's dirty little secret	<p>This is not a taboo subject but it's not something you talk about with everyone so the discussions opened up that avenue to be able to say to someone: how were your SETs? or I had this type of SET result, have you ever had that experience?</p> <p>I don't have an isolated experience, it's a shared experience. I think that that's something that, at a broader institutional level, is worth looking at because I would suspect it's not unique to our discipline.</p> <p>This is teaching's dirty little secret. We don't talk about this, we don't discuss this, there's no orientation around preparing for this, there's nothing</p> <p>None of us were prepared for this, none of us were informed that we were going to get unsolicited and unprofessional and uninformed feedback on our teaching that really impacts us personally</p>

<p>Feeling hurt and distressed</p>	<p>Comments that were...about me as a person... that were very hurtful...really had an impact on me</p> <p>I can lose sleep, have ruined weekends and have lingering thoughts about...some of the really hurtful negative comments</p> <p>I definitely notice things like ruminating, sleeplessness, being distracted, and second-guessing what I've done and the work I've put into a course</p> <p>You just kind of deal with it until the feelings numb and then you just wait for the next term and the next round</p> <p>It was like a punch in the stomach</p> <p>Negative comments are like a hidden bomb</p>	<p>It still hurts</p>	<p>I will say this time I was able to let it go much quicker than I have in the past</p> <p>I would say it was not a big deal, just a kind of unease</p> <p>I think regardless of how prepared you are, it is never 'nice' hearing negative things, especially if it isn't constructive criticism</p> <p>Physically - I think it's just that tightness in the chest that you get, very minor, but like, oh I need to be better or I didn't reach that student and that sucks</p> <p>It's just one student comment out of the whole group that sticks in your head</p>
<p>I perform, I act for SETs</p>	<p>I almost cater to the SETs</p> <p>I think I am mindful of the SETs all the time</p> <p>I'm often making decisions about what I'm going to use to evaluate students</p> <p>Maybe I would want to be a little more "you're an adult learner you could find the answer," instead I'll walk them through it because I am worried that they'll say I wasn't helpful</p>	<p>Still performing, still acting</p>	<p>I just think I feel a need to please all students...I seem to not be able to let go of that</p>

Table 4  
*Mean Student Responses for the Standard and Intervention Courses*

	Standard Course Mean <sup>1</sup>	Intervention Course Mean <sup>2</sup>
The textbook and/or assigned readings contributed strongly to this course	4.4	4.2
I found the course content challenging	4.1	3.9
I consider this course an important part of my academic experience	5.0	4.9
I would rate this course as very good	4.9	4.7
Students were treated respectfully	4.9	4.9
The instructor was available to students outside class	4.9	4.7
The instructor responded effectively to students' questions	4.9	4.7
The instructor demonstrated a broad knowledge of the subject	4.9	4.9
The instructor showed enthusiasm for the subject matter	4.9	4.9
The instructor encouraged student participation in class	4.9	4.7
The instructor set high expectations for students	4.7	4.4
The instructor fostered my interest in the subject matter	4.8	4.6
The instructor effectively communicated the course content	4.9	4.7
The instructor used class time effectively	4.9	4.7
Where appropriate, the instructor integrated research into the course material	4.2	4.2
The instructor provided effective feedback	4.7	4.4
Given the size of the class, assignments and tests were returned within a reasonable time	3.8	4.8
The evaluation procedures were fair	4.8	4.7
I would rate this instructor as very good	4.9	4.9

*Note.* <sup>1</sup>Response rate for standard course = 72% (45/59). <sup>2</sup>Response rate for intervention course = 68% (47/69), Rating scale ranged from 5 (*Strongly Agree*) to 1 (*Strongly Disagree*).

Table 5  
*The Nature and Tone of Student Comments*

	Standard Course	Intervention Course
Total # of “positive” comments	90	93
Total # of “negative” comments	26	35
Average # of words in “positive” comments	44	60
Average # of words in “negative” comments	28	30
# of comments containing absolute adverbs “always, never, none”	17	20
# of “negative” comments containing absolute adverbs “always, never, none”	0	1
# of comments containing adverbs “sometimes, usually, often, occasionally”	3	13
# of “negative” comments containing adverbs “sometimes, usually, often, occasionally”	2	12
# of comments containing positive metaphors	0	1
# of comments containing negative metaphors	0	0
# of comments about a specific problem	20	29
# of comments about an abstract issue	1	0
# of unclear comments (comment lacks clarity, difficult to understand)	0	0
# of comments focused on something that cannot be changed (curriculum rather than course, room location, type of seating etc.)	6	6
# of personal comments (focused on instructor’s body type, posture, voice, hair, clothing etc.)	1	1
# of negative personal comments (focused on instructor’s body type, posture, voice, hair, clothing etc.)	0	0
# of comments with recommendations	5	10
# of cynical comments	0	0
# of disrespectful comments	1	0
# of derogatory (belittling, hurtful) comments	0	0
# of acerbic comments	0	0
# of threatening comments	0	0

The “information rich case” instructor received a score of 4.9 out of 5.0 for the question “I would rate this instructor as very good.” This indicates that students in both sections regarded this instructor positively. We recognize that SET results from a less valued instructor might demonstrate a greater degree of differences in scores between the section that received the intervention and the section that received standard course content.

Mean student numerical responses remained relatively stable with the exception of “assignments and tests were returned within a reasonable time”. Students who received the in-class intervention had a much more favorable perception of the return time. Students who received the intervention provided marginally lower SET scores in thirteen questions than students who received standard course content. This might be because students who received the intervention were more thoughtful and used more discretion when answering the SET survey.

With regards to comments, students who received the in-class intervention submitted more comments about perceived course or instructor weaknesses (negative comments) compared to students who did not receive the intervention, but used more qualifiers such as “sometimes” rather than absolutes such as “always”, addressed specific issues, and made more recommendations for improvements: *Sometimes the content is delivered a little too quickly and it is hard to keep up with the notes, however she often addresses this problem by repeating what she said and answering questions.*

Their positive comments related to perceived strengths of the course or instructor were also more descriptive (more expansive) compared to comments from students who did not receive an intervention: *The instructor made sure to keep things exciting during class. She used a variety of materials from power points, to videos, to case studies to foster every learning style. Her power points were also to the point and made it easy to study. This really helped during exam time.*

Although they had received an intervention, students in both sections continued to comment about course or classroom features that cannot be changed such as room location, type of seating and/or the curriculum: *It was a long class so it was hard to focus for the full 3 hours.*

### **Quantitative Results (Instructors)**

The lunch-hour group discussions with instructors focused on providing peer support and time for contemplation of experiences receiving SET results. Instructors’ physical and emotional responses to SETs pre and post discussion groups are summarized in Table 6.

Prior to participation in repeated group discussions, instructors reported eating too much and drinking alcohol to cope with SET results. There were fewer reports of alcohol use and of over-eating post discussion groups. Although instructors reported that they continued to feel apprehensive, uneasy, and uncertain about impending SET results, they reported experiencing fewer physical symptoms such as fatigue, tense muscles, headaches and difficulties falling asleep, and less lingering feelings of sadness post SETs. Reports of reduced enthusiasm for teaching (related to negative SET results) also decreased post discussion groups.

To allow for comparison of qualitative and quantitative results obtained in a mixed methods study, Creswell and Plano-Clark (2018) recommend a table with congruent or incongruent findings on the horizontal dimension and the different themes, topics, or participant types on the vertical dimension (Table 7).

Table 6  
*Physical and Emotional Responses of Instructors to SET Results Pre and Post Discussion Groups*

	Pre-Intervention (n = 5)	Post Intervention (n = 4)
Headache	1 (20%)	0
Chest tightness	2 (40%)	3 (75%)
Rapid heart-rate or palpitations	2 (40%)	2 (50%)
Nausea	2 (40%)	1 (25%)
Shortness of breath	1 (20%)	0
Sweating	1 (20%)	0
Tense muscles, sore neck or back, or jaw pain	4 (80%)	2 (50%)
Fatigue	3 (60%)	1 (25%)
Difficulty falling asleep	4 (80%)	1 (25%)
Difficulty staying asleep	0	0
Loss of appetite	1 (20%)	2 (50%)
Eating too much	3 (60%)	0
Use of alcohol	3 (60%)	1 (25%)
Use of pharmacotherapy	0	0
Diarrhea, cramps, gas, constipation	0	0
Restlessness, itching, tic	1 (20%)	1 (25%)
Feeling of impending doom	1 (20%)	0
Feeling of uneasiness	3 (60%)	2 (50%)
Feeling of apprehension	3 (60%)	3 (75%)
Feeling uncertain	5 (100%)	4 (100%)
Feeling overwhelmed	3 (60%)	2 (50%)
Unable to concentrate	2 (40%)	1 (25%)
Irritability	2 (40%)	1 (25%)
Experiencing bouts of anger/hostility	0	1 (25%)
Feeling sad	3 (60%)	1 (25%)
Feeling hopeless	0	0
Loss of interest	1 (20%)	0
Lack of enthusiasm	4 (80%)	1 (25%)



Table 7  
*Joint Display of Qualitative and Quantitative Results*

Participants	Qualitative Results	Quantitative Results
Instructors	Pre Intervention	Pre Intervention Symptoms
Instructor perceptions, physical and emotional responses to SETs pre and post three lunch-hour discussion groups	Apprehensiveness	Physical symptoms included chest tightness, palpitations, nausea, tense muscles, difficulty falling asleep, eating too much, feeling sad, use of alcohol
	Uncertainty	Post Intervention Symptoms Fewer reports of tense muscles and difficulty falling asleep Fewer reports of sadness and irritability Reduced reports of alcohol consumption Some symptoms such as apprehensiveness, uneasiness, and feelings of uncertainty remained relatively stable
	Procrastination	
	Frustration	
	Unprepared <i>“It was like a punch in the stomach, a hidden bomb”</i>	
	Reduced enthusiasm for teaching	
	Pedagogical constraints <i>“I cater”</i>	
	A tendency to focus most on negative comments <i>“It lingers”</i>	
	Feeling tormented by negative comments	
	Post Intervention	
	Feeling prepared to receive results	
	Feeling less isolated <i>“I somehow wasn’t alone”</i>	
	Reduced procrastination	
New strategies for opening, reading, and interpreting SET results		
Changing what feels detrimental and debilitating	Post Intervention	
Making the most of constructive feedback	Fewer reports of lost enthusiasm for teaching	
Letting feelings of depression and frustration go <i>“I let it go”</i>		
Unable to stop catering		

Participants	Qualitative Results	Quantitative Results
<p data-bbox="199 272 317 300">Students</p> <p data-bbox="199 345 520 444">Student mean numerical responses and comments on SETs</p> <p data-bbox="199 490 520 737">Students who received standard course content compared to students who received an intervention in addition to standard course content</p>	<p data-bbox="562 272 953 300">No qualitative data to compare</p>	<p data-bbox="1350 272 1873 407">Student mean numerical responses remained stable with the exception of “assignments and tests were returned within a reasonable time”</p> <p data-bbox="1350 454 1873 630">Students who received the intervention submitted more negative comments but used more qualifiers, addressed specific issues, and made more recommendations for improvements.</p> <p data-bbox="1350 677 1873 776">Students who received the intervention were more descriptive (more expansive) in their positive comments</p> <p data-bbox="1350 823 1873 997">Despite the intervention, students continued to use the SETs as an avenue to complain about curriculum, course, or classroom characteristics that cannot be changed</p>

## Discussion and Implications for Practice

Our roles as researchers were to maintain neutrality while giving instructors a voice and our results are an interpretation of their experiences. Although most instructors who lecture in post-secondary educational institutions receive SETs, anecdotal evidence suggests that the majority feel anxious about receiving SET results and do not discuss their results with peers. Of those who do have a conversation with peers, most briefly mention their feelings of discomfort or distress and do not divulge details. Our results suggest that repeated group discussions that include time for reflection and peer support can have a positive impact on instructors' perceptions of SETs, reduce potentially harmful coping strategies, and shift basic approaches to opening, reading, and interpreting SET results. The participants also revealed that when they were hired by the post-secondary educational institution, they did not receive any formal preparation for the personal impact that SETs can have. Faculty departments, therefore, should commit to supporting SET discussion groups and create a formal orientation about the purpose and use of SETs for all instructors.

Participants described continuing to feel wounded, sad, and offended after reading negative feedback despite feeling supported through the discussion groups. Some symptoms such as apprehensiveness, uneasiness, and feelings of uncertainty did not change regardless of the intervention. These results suggest that, going forward, instructors should have a concrete strategy designed to help them critically review student comments. A tool for instructors, loosely based on the Table 5, should be developed to help them strategically and systematically review their SET results. Using this tool, instructors could categorize and tally comments as negative or positive, constructive or non-constructive, relevant or irrelevant while assigning a percentage to each occurrence. Instructors have previously suggested that they focus most on the negative feedback and negative comments can eclipse the positive comments. A quantitative tool could underscore positive and constructive feedback and assist previously distressed instructors to separate their personal feelings from their professional results with confidence.

Providing professional, thoughtful, mindful, and considerate feedback as well as understanding the potential negative influences of disingenuous or condescending feedback are essential skills for all post-secondary students. While students receive and deliver an abundance of feedback throughout their education, little or no time is allotted to teaching students how to deliver formal and informal feedback to others in a respectful and effective manner. In our nursing school, we have begun building a foundation for first-year students. In the fall of 2018, a select group of fourth year nursing students delivered the same 15-minute SET educational workshops to their peers in other years of the program as part of a leadership project. Peer teaching is known to have a positive influence on student knowledge acquisition (Brannagan et al., 2013) and peer led workshops can reduce perceived or potential biases related to having an instructor provide the same information. Our students also informally surveyed their peers after the workshops. The majority of student indicated surprise that the SET results were read and taken seriously by the instructors and the university, and were astonished to learn that SET results were used during tenure and promotion reviews. They also reported that they found the information helpful and that short workshops improved their confidence in providing constructive feedback to instructors. In the future, peer led workshops could be tailored to each year of an undergraduate program; first year students would receive foundational information and in later years the workshops would increase in complexity.

## Limitations and Future Research

While this study enhanced our understanding of instructors' physical and emotional responses to SET results and highlighted the potential benefits of peer led workshops for students and peer support discussion groups for instructors, there are some important limitations that should be acknowledged. First, the qualitative analyses are our interpretations of the stories that we heard during discussion groups and interviews. They do not necessarily represent the perceptions of other instructors or instructors from other disciplines or other universities. Similarly, the themes are ours and do not necessarily represent themes that would be identified by different researchers. While it was our intent to present an objective summary of the themes, the immersive nature of qualitative research makes it impossible for our personal biases to be completely separated from this work. Despite our structured efforts to maintain rigor, our biases may have affected how our results were generated during the discussion groups, interviews and data analysis. Second, the sample size was small and all instructors who participated were teaching in a school of nursing. While the results may be transferrable to other disciplines, they are not generalizable. The same methodology should be repeated in other faculties, at other universities, and with larger sample sizes. Finally, in order to avoid skewing the SET results the students were not informed about the study or about the real reasons for the three 15-minute workshops. Unfortunately, this step prevented any formal data gathering from students pre and post workshops. In a future study students could be surveyed both qualitatively and quantitatively pre and post workshops to determine what they had learned and how what they had learned might change their feedback practices. One survey question could be "If you could go back in time, would you change any of your previously submitted SET comments? If yes, why?"

## References

- Beleche, T., Fairris, D. & Marks, M. (2012). Do course evaluations truly reflect student learning? Evidence from an objectively graded post-test. *Economics of Education Review*, 31(5): 709-719. <https://doi.org/10.1016/j.econedurev.2012.05.001>
- Beran, T. N. & Rokosh, J. L. (2009). Instructors' perspective on the utility of student ratings of instruction. *Instructional Science*, 37, 171-184. <https://doi.org/10.1007/s11251-007-9045-2>
- Brannagan, K. B., Dellinger, A., Thomas, J., Mitchell, D., Lewis-Trabeaux, S., & Dupre, S. (2013). Impact of peer teaching on nursing students: perceptions of learning environment, self-efficacy, and knowledge. *Nurse Education Today*, 33(11), 1440-1447. <https://doi.org/10.1016/j.nedt.2012.11.018>
- Chan, K. (2001). The difficulties and dilemma of constructing a model for teacher evaluation in higher education. *Higher Education Management*, 13(1): 93-111.
- Chen, Y & Hoshower, L. B. (2003). Student evaluations of teaching effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education*, 28(1), 71-88. <https://doi.org/10.1080/0260293032000033071>
- Creswell, J. W., and Clark, V.L.P. (2018). *Designing and conducting mixed methods research* (6th ed.). Thousand Oaks, CA: Sage.
- Darwin, S. (2016). *Student evaluation in higher education: Reconceptualising the student voice*. New York, NY: Springer. <https://doi.org/10.1007/978-3-319-41893-3>

- Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., & Van Petegem, P. (2013). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Education*, 36, 1-11.  
<https://doi.org/10.1016/j.tate.2013.06.011>
- Farr, M. (2018, August 28). Arbitration decision on student evaluations of teaching applauded by faculty. *University Affairs*. Retrieved from <https://www.universityaffairs.ca/news/news-article/arbitration-decision-on-student-evaluations-of-teaching-applauded-by-faculty/>
- Gravestock, P., & Gregor-Greenleaf, E. (2008). *Student course evaluations: Research, models and trends*. Toronto: Higher Education Quality Council of Ontario. Retrieved from <http://www.heqco.ca/SiteCollectionDocuments/Student%20Course%20Evaluations.pdf>
- Moore, S., & Kuol, N. (2005). Students evaluating teachers: Exploring the importance of faculty reaction to feedback on teaching. *Teaching in Higher Education*, 10(1), 57-73.  
<https://doi.org/10.1080/1356251052000305534>
- Roberts, R.C. (2015). Are some of the things faculty do to maximize their student evaluations of teachers scores ethical? *Journal of Academic Ethics*, 14, 133-148.  
<https://doi.org/10.1007/s10805-015-9247-1>
- 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.  
<https://doi.org/10.3102/0034654313496870>
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research*. Thousand Oaks, CA: Sage.
- Sunderji, N., & Waddell, A. (2018). Mixed methods convergent study designs in health professions education research: Toward meaningful integration of qualitative and quantitative data. *Academic Medicine*, 93(7), 1093.  
<https://doi.org/10.1097/ACM.0000000000002241>
- Thomas, D.R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237-246.  
<https://doi.org/10.1177/1098214005283748>
- Uttl, B, White, C. A., & Wong Gonzales, D. (2016). Meta-analysis of faculty's teaching effectiveness. Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, 54, 22-42.  
<https://doi.org/10.1016/j.stueduc.2016.08.007>
- Winer, L., DiGenova, L., & Costopoulos, A. (2016). Addressing Common Concerns about Online Student Ratings of Instruction: A Research-Informed Approach. *Canadian Journal of Higher Education*, 46(4), 115-131.
- Yao, Y., & Grady, M. L. (2005). How do faculty make formative use of student evaluation feedback?: A multiple case study. *Journal of Personnel Evaluation in Education*, 18(2), 107-126. <https://doi.org/10.1007/s11092-006-9000-9>