Focus Group Analysis: A Summary of Key Themes from Students in all Faculties

The Course Evaluation Project Team (CEPT2)

September, 25 2019

Sonya Buffone

Table of Contents

Contents

Focus Group Analysis: A Summary of Key Themes

Introduction

CEPT(2)'s remit from Senate involves timely development of recommendations for the roll-out of a specific process for gathering student perceptions of teaching in all parts of campus. In particular, CEPT(2)'s remit includes:

a) Recommendations with respect to a specific, cascaded questionnaire that includes a set of core questions for use in all Faculties (and AFIWs that choose to participate)
b) Development of "Users' Guides" for instructors, Chairs, Performance Review Committees, and Tenure and Promotion Committees (TPCs), which will advise on the appropriate interpretation of results
c) Recommendations about process, software to be used, and related issues

CEPT(2) intends to build as much as possible on the good work of CEPT(1). One thing CEPT(1) developed was a draft set of core questions. CEPT(2) decided to hold a series of focus groups with undergraduate students to inform its decision about what modifications, if any, to make to the draft instrument.

Once the decision was made to hold focus groups, we also took the opportunity to investigate other matters for which the focus group format fit. In particular, the research literature on SETs suggests that the following areas are under-investigated: what students understand themselves to be doing when they complete end of course surveys, and what they understand good teaching and a well-taught course to be in the context of such surveys (see for example, Gravestock & Gregor-Greenleaf, 2008; McKeachie, 1997; Ory & Ryan, 2001; Willits & Brennan, 2017). Further, previous research suggests that there is significant overlap between student/faculty perceptions about what constitutes quality teaching (Wachtel, 1998; Ory, 2001). Focus groups with Waterloo undergraduate students seemed especially likely to provide useful data about how students perceive effective teaching and to inform the Users' Guides.

Overall, the following three research aims guided the data collection process:

1. Determine how students understand/perceive a positive/negative course experience and examine how well these perceptions are aligned with evidenced-based principles of effective teaching as proposed in the University of Waterloo's Undergraduate Learning White Paper Report and how well these views align with the core questions proposed by CEPT1;[1]
2. Examine how students interpret and understand items from the draft set of core questions and;

---

[1] These principles are drawn from the literature. While the ideas conveyed in the report have not been formally endorsed by the university (e.g. by a motion at Senate), at the time of the focus groups the White Paper had been produced by a working group as part of the strategic planning process focused on teaching and learning. The group included various campus stakeholders (faculty, administrators, students and staff), and was available to the campus community during the strategic planning process.

3. Understand the response processes of students:
    a. What are the underlying motivations for completing course evaluations (e.g.., are they more likely to complete surveys if they hold extreme opinions?), and how do students think course evaluations are used (i.e., formative/summative)?
    b. How do students use the five-point Likert scale as proposed by CEPT(1)? (e.g., what does selection of three signify? And how do they decide between the upper and lower extreme ends of the scale?)

Brief Summary of Methodology

The methods used for data collection and analysis are outlined in this document in brief. A full discussion of these procedures is provided in the Methodology Appendix (see https://uwaterloo.ca/waterloo-course-evaluations/focus-group-methodology).

In total, six focus groups were conducted, one for each of the six Faculties.

- Students were recruited primarily via social media and mass email communications (e.g., through the Registrars Office).
- Participants were required to register to participate via an online form.
- We received 259 requests to participate, which is quite substantial given the timing (Spring term).
- Invitations aimed to achieve representativeness of the student body with respect to gender, year of study, diversity of programs, etc., in each group.
- Each session was approximately 90 minutes, and all of the focus groups were held in July 2018.
- Each session was audio recorded, and then transcribed (verbatim) by an undergraduate research assistant.
- The transcripts were then analyzed by the Senior Research Specialist for CEPT(2).

Conducting the Focus Groups

At each focus group, students were given two documents (see https://uwaterloo.ca/waterloo-course-evaluations/focus-group-discussion-guide). Document 1 included the questions used to guide the group discussion for the first portion of the session. Students received document-one when they arrived at the focus group. Document 2 included a copy of the draft set of core questions proposed by CEPT(1). It is important to note, students **did not receive** document 2 until we had completed our discussion of the items on document 1 (which usually occurred at about the midpoint of the session). The decision to separate these documents was intentional. We did not want to sensitize students to any of the items used in the draft set of core questions, as we recognized the potential for this to influence how students responded to the following questions (from document 1): Tell me why you would give a low evaluation, and; What are the qualities that make for a really great course experience? In this sense, we can have more confidence in the

authenticity of student responses to these questions given that they were not primed to consider specific items from the draft instrument in formulating their responses.

Data Analysis

The data were analyzed in two stages. The first phase of analysis involved inductive coding methods (Strauss & Corbin, 1998; Thomas, 2006). Each interview transcript was reviewed and codes were derived inductively by identifying patterns and consistencies as they emerged in the data, which led to the development of key themes. In qualitative research, the aim of the analysis is to achieve depth rather than breadth; data collection continues until a point of saturation is reached (i.e., no new themes emerge/information is redundant) (Glaser & Strauss, 1967). In focus group research, the general rule of thumb is to conduct three or four focus groups and then decide if additional groups or cases are necessary to include in the study (Krueger & Casey, 2009). In light of these considerations, we decided it was necessary to include all six Faculties in the sample, at the outset, to capture any potential variability in student responses.[2] As described in the summary findings below, across all six Faculties, we found considerable consistency and overlap in the key themes that emerged from the data.

The second phase of data analysis involved deductive coding methods (Strauss & Corbin, 1998). Codes derived from the interview data were compared and contrasted with four of the five evidence-based principles of effective teaching as proposed in the University of Waterloo's Undergraduate Learning White Paper Report (Alignment, Motivation, Inclusivity, and Deep Learning). This analysis was done to identify any potential relationships or alignment between the emergent themes (from phase one of the analysis) and the proposed key principles of teaching quality. In this sense, deductive coding methods provided the opportunity to address the second research aim of project.

In the quotations used throughout this document, students are not identified (to protect confidentiality), but are instead referenced by their preferred gender identity and an assigned number within each group (female = F; male = M &; other = O).[3] Additionally, the Faculties have been designated as follows: Arts = AR; Science = SC; Math = MA; AHS = AH; Engineering = EG and finally, Environment = EV. To provide an example, the first female student to the left of the moderator in the AHS focus group would appear as AHF1, and each time AHF1 appears in text it refers to the same participant.

Summary Results Overview

In this report, we provide a summary of the results from the six focus groups. We summarize each discussion topic thematically in the order in which the topics were raised, in each focus

---

[2] We anticipated variability in responses would be more likely across Faculties, as opposed to within Faculties.

[3] For all of the focus groups, numbers were assigned based on seat order. For example, the first female to the left of the moderator was designated as F1. The next female would be F2, F3, etc. The same logic was applied to males and those who identified as 'other'.

group. To illustrate each of the key themes we have identified, we provide an example, verbatim, of student responses.

Part one, discusses key themes that emerged with respect to students' understanding of the purpose and uses of course evaluation data. Parts two and three provide a summary of the key themes as they relate to students' perceptions of what defines a poor/excellent course experience. At the conclusion of parts two and three, we provide a summary of how the key themes identified in this analysis relate to the evidence-based principles of teaching effectiveness as proposed in the University of Waterloo's Undergraduate Learning White Paper Report, as well as how the themes relate to the SCP tool proposed by CEPT(1). Part four discusses students' perceptions about the five-point Likert scale, and finally part five discusses students' interpretations of three items from the draft set of core questions. In each part of the document, we also describe some recommendations that might be drawn with respect to the mandate of CEPT(2).

Part 1: Student Understandings of Purpose/Uses of Course Evaluation Data

A. Purpose of Course Evaluations

Guiding Question: Why do you think students are asked to complete course evaluations?

In response to this question, students overwhelmingly assume that course evaluations are used for formative purposes, though some students recognize (or at least speculate) they are also used for summative purposes.

Theme	Description	Example
Formative	Perception that course evaluations are used to improve instructional practices, or the course in general.	I am just thinking more along the lines of maybe sometimes they were just seeking feedback. Like, maybe they were trying something new. So not even "is my teaching good or bad", but maybe, "I am experimenting with a new teaching method," or something like that (MA02).
Summative	Perception that course evaluations are used for summative purposes (e.g., promotion & tenure reviews).	F2: Um, I'm not sure if it is true or not, but sometimes we do course evaluations and they affect tenure position. Um yeah, they affect tenure positions (ARF2).

Resulting Recommendations

- Create toolkit (or web platform) for students to access to:
  - Explicitly inform students that these tools are used both formatively and summatively.
  - Clarify the importance of these tools for teaching at UW since students lack a clear understanding of why we use SCPs (e.g., why do we value them? Why should students complete them?)
    - Instructors can also play a role in this process (e.g., explain why students in their course(s) should participate in course evaluations; describe changes that have been made in response to past evaluations, etc.)

B. Why submit a course evaluation?
Guiding Question: Explain why you would complete, or choose not to complete, a course evaluation at the end of the term.

Overwhelmingly, the majority of students indicated that completion is more likely when **feelings are extreme** (high or low ratings), but other reasons motivating completion were also outlined. The table below summarizes these perspectives.

Theme	Description	Example
Extreme Feelings	Desire/willingness to complete course evaluations resulting from extreme like/dislike of the instruction in the course/overall course experience.	Yeah, I am more inclined to do a course evaluation where the course went really, really well and I liked everything about it. Or if there are problems in the course, or if I was frustrated at some point. But if a course just okay then I won't necessarily want to do the evaluation (EVF2).
Helping Future Students	Desire/willingness to complete course evaluations (regardless of positive/negative experience) to help future students (the underlying assumption here is that feedback will be used to make improvements).	I personally just feel that it is important for them to have feedback to go off of whatever the student thought, for future courses. Just because I feel like if they never get feedback then they will never know, right? So, I personally just complete every course evaluation, just because I feel motivated to do it because then I feel like it might help future students (SCF1).
Helping Instructors	Desire/willingness to complete course evaluations to help instructors (e.g., provide feedback to improve course or instructional methods).	I never skip evaluations because I do believe it does help profs do better next time. So, if I don't like something I say what I don't like. But then I also do a suggestion. Instead of just saying, 'no

		that's bad,' I say, 'Maybe if you could do this maybe it would be better for the class,' or whatever. So yeah, when it is good I say, 'Thank you for doing this, this, and this and this was awesome.' So, kind of appreciating and saying, 'Thank you, thank you, thank you.' (ARF4)
Instructor Attitude	Desire/willingness to complete course evaluations resulting from the impression that the instructor sees the activity as valuable/worthwhile.	So, typically, I don't give low evaluations for teaching because I think the profs that do deserve it just don't care. So, I have the attitude of, I don't care either, so I might not choose to do an evaluation for them (ARF2).
Frustration with lack of long-term understanding (regarding whether data being used long-term?)	Desire/willingness to complete course evaluations is deterred because students feel frustrated that they will not see or experience any changes (based on the SCP results).	Definitely, what I would like to see in evaluations is even just feedback, like are they doing something about it, because we are just submitting these things but we don't know what happens to it, or if they even look at it. Like, some profs are, like, yeah, they really want the feedback, so I can totally see that they might take it into future consideration for their growth and the growth of the course. But with some, I have no idea where it is going. So, it is like, 'Do I really need to submit an evaluation for that course?' (EVF3)

Resulting Recommendations

- In our instructional tools for Chairs and instructors, we should make it clear that students with extreme opinions are more likely to submit a course evaluation, but this goes both ways: those who enjoy the course and those who do not. This situation is also evidenced in the literature and is why we tend to see skewed distributions, **most often with scores falling at the high end of the scale between 3-5.**
- These results also motivate the earlier recommendation to ensure that students know that these evaluations are used both formatively and summatively, and so are important.
- It should be strongly recommended that instructors provide class time to complete evaluations. This is one way to show students that instructors care about student participation in SCPs. Instructional tools for students should highlight the importance of SCP surveys and explain clearly what they are used for. It might also be useful to

encourage instructors to highlight/explain changes that have been implemented as a result of previous evaluations.

- Encouraging instructors to provide mid-term opportunities for students to provide feedback, independently from the formal SCP process, would be another strategy for conveying that instructors value student feedback. Instructors should also share with students how previous course evaluations have been used to inform course design and activities.

Part 2: Low Evaluation of Teaching/Poor Course Experience

Guiding Question: Tell me why you would give a low evaluation of teaching?

Across all six Faculties, students overwhelmingly focused on the same fundamental reasons for assigning low evaluation scores. In fact, there was enough consistency across all six groups to generate two overarching themes (Negative Social Atmosphere & Poor Course delivery).

Negative social atmosphere	The first theme focuses on negative feelings about the instructor's willingness to help students, which creates a negative social atmosphere in the learning environment (Willits et al., 2013). This consists of two-subthemes: unapproachable/unavailable and disinterested.	
Subtheme	**Description**	**Example**
Unapproachable/Unavailable Instructor	Students feel unable to seek help, ask questions, etc. (or are uncomfortable doing so), or students perceive the instructor to be unavailable with respect to holding office hours or answering questions in/or outside of class.	They say. 'C'mon this is so easy/' Like, two profs have said that, like, 'C'mon, this is easy, like. this is easy, a piece of cake, what could you not understand?' And everyone is just looking around, like, 'Why?' This is, like, 'Why do they think this is so easy?' And that just makes them less approachable, too, because if I am not understanding from what you are teaching and you are expecting me to, just, to somehow just understand this, then, like. I felt like they felt that we need to understand it on their level. And that's what I felt with the two experiences that I had, and that's what turned me away from asking for help as well, unless I really could not figure it out on my own. It was just, like, 'I really don't want to approach this.' (EVF4)

Disinterested Instructor	Perception that the instructor does not want to teach the class/ doesn't enjoy teaching. According to students, some examples of behaviour that might indicate the instructor is disinterested include: monotone speaking, reading directly off their slides, etc.	For me, I would give a low evaluation because the professor does the bare minimum. Like, 'Yeah, I am here,' read off the slides, 'Okay, see you, you can email me.' (SCF3)
Poor Course Delivery	The second theme centered on feelings about the instructor's ability to communicate course content and expectations in a clear and understandable manner. This theme consisted of two-subthemes: Poor Communication and Unclear Expectations.	
Subtheme	**Description**	**Example**
Poor Communication	Overall perception that the instructor failed to convey information in a clear/understandable manner	The quality of the instruction has a lot more to do with how clear the information is getting across. Where if the instructor goes on this great tangent road and never ties it back in with what you're learning, as entertaining as that may be, it really can take away from what you are learning. And in those cases where either the information is not getting across, or the information is just so diluted by tangents, I might be more prone to give a low evaluation (ARM1).
Unclear expectations or assessments (Course Design)	Perception that the instructor has not clearly conveyed broader understandings of course concepts as evidenced by graded assessments (i.e., testing random facts, or things not taught). The biggest sentiment that emerged here was that in these instances,	For me it really largely depends on what the professor chooses to test you on. The content that they do test you on, to me, that's what they are saying is the most important in this course: this is what will help you in

	students expressed that they did not understand what was expected of them.	your studies. So if they just test you mostly on really small details and things that just don't really have that much to do with the course, I'm less inclined to give a good rating of that teacher because they're not really testing you on relevant material (ARM3).

Unique Cases

The tables highlighted general themes that emerged from the data and the associated subthemes that fall within each general theme. There were other responses that seemed to reflect the general themes, but did not fall squarely into the subthemes. These types of responses appeared less often in the data, but are worth noting given that they reflect the prominence of the general themes and the wide range of issues within these themes. Some examples are provided below:

> Reasons I could give them a low rating… because sometimes honestly thick accents, my ability to take in the information from the course (ARF2).

> - This issue reflects one aspect of communication clarity. The CTE provides advice and strategies to assist instructors with accents.

> If the prof or lecturer comes off as a little bit intimidating or in a sense rude. Kind of just have one example to give but there were students using their phones near the front so it is understandable why a prof would get agitated, it is a disrespectful thing to do, but this was back in first year when students have a different study or lecture approach than they do now and so the way the prof responded was very intimidating to the students and everyone was really shocked in how he responded. So, I would give a prof a low evaluation if they have sort of trouble dealing with students of that kind or sort of controlling their emotions in some type of situation, and it is understandable, it is not easy to do, especially when it's something like that, however because that could be a make or break point for that student and make them not want to come anymore or like even though they are still like adjusting at this point, it really is just like a first year issue it wouldn't happen more so in upper years, but if they can't sort of teach the students the importance of paying attention rather than just blatantly yelling at them or pointing them out in class and sort of like embarrassing them in front of the class (AHM1).

> - This example contributes to a 'negative social atmosphere'

> Then again, I had a quantum course where the average was like 50 or something and that shapes how I will fill out the evaluation like if the prof is just like failing half the class

then he is doing something wrong it is not the class that is doing something wrong (MAO1).

- This could be considered to relate to unrealistic expectations/inappropriate assessments.

Part 3: High Evaluation of Teaching

Guiding Question: What are the qualities that make for a really great course experience?

When asked to explain what factors contribute to a positive course experience, students discussed many of the same types of factors emphasized in their view of low evaluations.

Overwhelmingly, students across all six Faculties described the same fundamental considerations, which are captured by the following three themes: Course Delivery (instructor is a clear communicator, and clearly defines expectations); the instructor Engages Students (encourages participation, applies course content, and cares about course content) and finally the instructor fosters a Positive Social Atmosphere (is open to questions, and is accessible/available).

Course Delivery	The first theme centered on feelings about the instructor's ability to communicate course content and expectations in a clear and understandable manner. This theme consisted of two subthemes: Instructor is a Clear Communicator and Clear Expectations.	
Subtheme	**Description**	**Example**
Instructor is a clear communicator	Perception that course content was conveyed clearly and understandably and connections (i.e., between key course concepts/theories) were made clearly.	So, I think that something that really helps is that if they are really clear and structured in their teaching and they have a lot of examples that they can go over (EGM3).
Clear expectations (course design)	Perception that the instructor clearly conveyed expectations for learning (often students referred to feeling that they knew what was expected of them with respect to graded assignments).	Like, one course in particular, it was very clear from the beginning that the prof already knew what was going to happen through all of the lectures and through the whole course. And I really appreciated that as well, because I knew that someone knew exactly what was going to happen in that course (MAM3).
Instructor Engages Students	Perception that the instructor makes an effort to actively include students in the learning process.	
Subtheme	**Description**	**Example**
Encourages Participation	Students express positive response to instructors who attempt to include some form of open discussions during class	This particular instructor, I thought was particularly motivating. He really encouraged, tried to

	(e.g., asking students to actively participate in discussion during class time).	encourage, participation. Of course, as you all know, doesn't always happen. Um, he accepted all answers, you always felt that no matter what you said he would use it in a positive way, and then direct it in the direction he was hoping for (ARF1).
Applies Course Materials	Perception that applying course concepts to real world examples or experiences helps to foster understanding (e.g., including guest lecturers from industry, research articles, research experiences, linking material to news stories, etc.)	Also, a professor that can give, like, real life examples. I had one course where the professor was working in the field that he was teaching and he gave us some examples, like a certain health policy where he actually implemented it, and how that worked out, and who was affected or not affected, what are the benefits, what are the disadvantages, and things like that. So, we as students can figure out if we want to go into this field or do we want to switch to something else. Are we interested in it? Do we not like it? Just things like that. (AHF3)
Cares about course content (enthusiastic/engaging).	Perceptions concerning the instructor's (perceived) effort(s) to make the class interesting or "worth attending".	For me it's, like, if the course itself isn't something that I'm interested in, like, for example, I'm in statistics—the professor that is teaching, you can tell how much the professor likes what he is teaching, and you can tell how engaged and how hard he tries (AHF6).
Positive Social Atmosphere	Emphasis is on instructors who show they have a genuine interest in their students and make an effort to show they care about their students.	

Subtheme	Description	Example
Open to questions	Includes any reference about the instructor creating an environment where students feel comfortable asking questions/and or students feel that their questions are taken seriously.	When they are more warm, and you feel like you can go up to them and ask them questions, and they are there to help you. And, I know that this is an Environment focus group, but I am a Don at St. Paul's, and we have a lot of first year engineering students. And they are always complaining about how they didn't understand this thing in class, and I will say, "Why didn't you go ask your teacher?" And they say, "because we are scared of our teacher and we don't want to go ask them." […] In my experience, in the Environment Faculty teachers are always there to help you. You can go to extended office hours, you can email them, and they will explain it to you. Or you can go to the TAs, and the TAs are so helpful so for me. That contrast was just pretty baffling to me (EVF1).
Accessibility/availability	Includes any reference to the instructor making him/herself available to students inside/outside of class (e.g., office hours, email) for additional help.	I would consider how often they address the class, like, questions and emails too. Like, when they are not lecturing, do they just disappear? Is it possible to get ahold of them? (SCM1).

Summary of Findings from Parts 2 & 3

In general, findings from this analysis reveal three key things: First, there is a considerable degree of consistency in the factors students use to describe their decision to provide a high/low evaluation within and across Faculties. Second, consistent with the literature, students at UW not only view instructional quality as multidimensional (see Ory & Ryan, 2001), but they also tend to emphasize aspects of instructional effectiveness that have been cited extensively in the

literature on this topic (see Allin & Jopling, 2009 for a summary of this literature).[4] Third, ***the factors that students tend to discuss*** in reflecting on their course experience(s) appear to be well-aligned **with the evidence-based *principles of quality teaching* proposed in the University of Waterloo's Undergraduate Learning White Paper Report.**[5] In other words, students seem to emphasize the same key principles of their learning experience that appear in the white paper report.[6] The following table provides a summary of the conceptual linkages found between the key themes identified in the focus groups, the three guiding dimensions, the proposed evidence-based principles of teaching quality and specific items on the draft set of core questions.[7]

Table 1: Conceptual Linkages: Key Themes and Principles

Focus Group Sub-Theme	SCP Dimension	Proposed Evidenced-Based Principles of Teaching Effectiveness	Description of Link	Proposed SCP Item(s)
Clear Expectations	Course Design	Alignment	For students, alignment in design principles appears to be an important consideration. One of the most commonly (and consistently) cited reasons for assigning a low evaluation of teaching is when students feel the instructor has not clearly conveyed expectations about what they are expected to learn and what	I knew what I was expected to learn in this course. The graded work assessed what I was expected to learn. The course activities prepared me for the graded work.

[4] It is worth noting, based on an extensive summary of the literature that looks at instructional quality, Allin, Clarke & Jopling (2009) identify the following four dimensions: Supportive learning environment, Academic expectations, Scaffolding Learning and Clarity.

[5] As noted in footnote 1, the White Paper was produced as part of the University's 2020-2025 strategic planning process by a stakeholder committee. In the absence of a Senate endorses statement of teaching principles, CEPT(2) took it to be a good proxy for opinions in the campus community about the nature of effective teaching.

[6] .Consultative review is ongoing at this time.

[7] As proposed by CEPT(1), in the draft report of the core questions for the new student course perception survey, the three guiding dimensions include: Course Design, Course Delivery & Learning Experience.

			assignments/tests will assess.	
Feedback for Improvement	Course Delivery	Motivation	Students clearly value feedback that provides the opportunity to improve upon their work. This was the single most important factor students discussed with respect to this item.	The instructor returned graded work in a reasonable amount of time.
Clear Communication	Course Delivery	Deep Learning	Students emphasize that they value an instructor who conveys course content clearly and understandably.	The instructor was a clear communicator.
Open to questions *Accessible/available*	Course Delivery	Inclusivity	In their discussion of low/high evaluations, students consistently and repeatedly highlighted that they value instructors who create a 'positive social atmosphere'. In this context, students felt comfortable asking questions and also expressed feeling that the instructor was willing (and available) to help students, thereby fostering a supportive learning environment for students.	The instructor created a supportive environment that helped me learn.
Encourages participation *Applies course materials*	Course Delivery	Deep Learning	Students indicated that they tend to rate instructors higher who try to stimulate interest in the	The instructor stimulated my interest in this course.

			course. It is clear that students value instructors who make an effort to engage them, by encouraging participation, applying course concepts to real world examples and demonstrating their own interest in the course.	
Cares about course content				

It is important to note that three items included on the proposed SCP instrument did not appear to align with any of the themes identified in the focus groups, including:

1. The course workload demands were…
   - o With respect to this item, students did not appear to connect workload with quality teaching practices. These findings are interesting given that there has been considerable debate in the literature concerning perceived workload and SET scores. Some researchers have argued that low-workload courses receive higher SET scores (Greenwald & Gilmore, 1997), while others (Marsh, 1982) find an inverse relationship where courses with more work receive higher SET scores. Similarly, Marsh & Roche (2000) found SETs were not biased by low workload, while Centra (2003) found a non-linear relationship between workload and SET ratings. Finally, in their study that looked at 500,000 classes across more than 300 institutions, Benton et al. (2013) found students rate instructors higher when they feel challenged. In light of such inconsistent findings, it is, at the very least, worth noting that students did not mention workload as an important consideration when evaluating instructors.

2. Overall, I learned a great deal from this instructor.
3. Overall, the quality of my learning experience in this course was excellent.

There is little consensus in the literature concerning the use of 'overall/global' ratings of instruction. Some researchers support the use of overall ratings (see Abrami, & d'Apollonia, & Rosenfield, 1997). On the other hand, one of the principal critiques of such measures is that they fail to adequately represent the multidimensionality of teaching (Marsh, 2007). Moreover, some research suggests that such measures are more prone to bias (Benton & Cashin, 2014; Cohen, 1990; Theall & Franklin, 1990). This lack of agreement highlights the importance of assessing how these particular items perform on our proposed SCP instrument during pilot testing.

A further point worth noting is that some students expressed some awareness that factors other than the quality of instruction might influence their ratings, some even expressed varying levels of desire to compensate for such factors. For instance:

> Another thing too would be I feel like the students know which courses they want to take and that also may have an effect on that. Like, I am not really a big fan of marketing, and I really don't think I could possibly take — Like, I am only taking it because it is required for my degree — but I don't think I could take a marketing course where the prof could be so good — like, it could be the reason that I don't like marketing, because I have always had bad profs, but it could be that I don't like the course material, and as a result like the prof can't really make me excited about the material just because I do not like the material and as a result I may see the prof as being poor (MAM1).

> A reason for maybe not filling out the course evaluation would be because I have a bias towards the course could be that I am not very good at, so I think that just because there are hard tests that is not necessarily the professor's fault. It's just, it's not just my way of thinking and so I don't want my ratings of the professor, could be he did a really good job, but I might not really see that because [in that] course I didn't really have a great time (ARM3).

> I also feel that with course evaluations to make them more like effective, I don't know what a good solution would be, but some way to account for the student also taking responsibility because I know there was one course that I found that it wasn't the prof's fault. It was a bland course and he tried his best to make connections and when he did make connections to the real world like about business or whatever it is like I listened, but I also didn't put in the effort and I would zone out a lot. So, I don't know, like I can recognize that and like if I was giving him a rating on how he was as a prof I wouldn't give him a low rating I would give him a higher rating. I can recognize when it is me, but I am not sure that everyone can make that distinction especially if you get a lower grade it is easier, you get more frustrated and you are more inclined to want to blame the prof like I didn't understand it because of them whatever (EVF4).

Though it is encouraging to find some students appear to recognize how certain biases might impact their ratings, we cannot assume that all students possess this same level of self-awareness. Therefore, we should look for ways to ensure this information is conveyed clearly in our instructional tools.

Implications/Recommendations

- We need to ensure that our educational tools for students, instructors, and Chairs highlight and explain the evidence-based principles of effective teaching. It is important to emphasize that our SCP tool aims to measure these key principles. Students should be aware of these elements when completing the SCP survey.

- The importance of the links found between themes (the focus group) and the proposed principles of teaching effectiveness, outlined in Waterloo's Undergraduate Learning White Paper Report, should not be understated. This finding suggests that **aspects of a course experience that students tend to value are well aligned with evidence-based principles of effective teaching as articulated in the White Paper, and SCP items identified by CEPT1.** This means that the questions proposed by CEPT(1) seem appropriate to use to measure factors related to learning experience(s) that students also consider when assessing their experience in a course. Interesting to note is the fact that the themes that emerged in this study are also highly consistent with literature that has found students are in the best position to evaluate certain aspects of their course experience (see for example, Nasser & Fresko, 2002; Theall & Franklin, 2001; Ory & Ryan, 2001; Wachtel,
- 1998; Wagenaar, 1995).

Part 4: Scale Understandings

Focus Group Guiding Questions

Students were provided with a blank Likert scale with 5-response categories (1= strongly disagree & 5= strongly agree). They were given time to examine the scale and then they were asked to following questions:

1. When you see questions that ask you to respond on a '5-point scale about agreement' what does the middle number '3' on the scale mean to you? (E.g., if you select '3' what does that mean? Does it reflect your reluctance to respond to the item?; do you feel it genuinely reflects a "middle" ranking?; do you select this when you are unable to make a decision about how to rate the item?)
2. Decision-making Processes:
   a. On this type of scale, what does a "low" evaluation look like?
   b. On this type of scale what does a "high" evaluation look like?
   c. On this type of scale, what would you select if you didn't have an answer/didn't know the answer to the question asked?

Use of '3'

When the 3 on a five-point Likert scale is not clearly labelled, there is significant confusion about its meaning. This is problematic because it becomes the 'go-to' response for several perceptions (e.g., "avg. rating of instruction", "don't know", "not applicable", "no opinion", etc.):

> If I am indifferent about it. Like, if it is something that is just so-so. Like, if I am neutral about it, like — it didn't affect me either way… the super positive or super negative —

then the three is, "I don't have an opinion and it is not my place to say yes or no" (AHF6).

> When you don't really have anything to say about a course […] what I would do is skip all the comment questions and for the scale questions just put 3 everywhere, so I am giving some data instead of just forgetting to fill out the evaluation (MAM2).

> A 3 is like, "They did a lot of things good, but could be improved upon," or maybe just, "The material itself was dry, and sometimes there is just no way to improve on that it." Just, "wasn't a bad course, but it was just ok, it wasn't amazing but it was ok" (SCF4).

Resulting Recommendations:

- Clearly label all points on the scale
- Include "no response"/ "prefer not to answer"/ "not applicable" option, a comment box to explain this selection may also be valuable.
- Provide students with a clear explanation of how to interpret the scale appropriately.

Decision-making processes (e.g., decisions about which point to select on the scale)

Many students suggested that they make intentional distinctions between different values on the scale (i.e., 4 or 5, 1 or 2), but there is little consistency between the distinctions they say they are making. For example, the decision-making process is simplistic for some, and complicated for others, and sometimes it seems logical, but other times it does not.

> Some people say, 'I gave you a 4 but that is the highest I will ever give!' But then it's like, it should be a 4-point scale, not a 5-point scale. Yeah, like, why would you do that if that's the best? So, yeah, 5 to me is a 5. And I would use 1 very infrequently, because most people put some sort of effort into things and I don't want to discourage them. It's just if they need some tweaking 2 or 3 comes in (AHF1).

> I mean, yeah, 5 is definitely excelling at everything, like, "I really loved the course, everything was great, and the instructor was great, so whatever they were doing they should continue doing." 4 is more like, "it was ok, but I can't really think of a good solution, but there is definitely still room for improvement." 3 is like, as I talked about earlier, is like, satisfactory so like it meets requirements, but again I can't really, I probably would write a lot more comments on it. 2 and 1 I could hardly ever give, like 1 is like I strongly hated the course. But knowing my personality I would've told the instructor a long time ago, directly, before even waiting for the final evaluation, because at the evaluation it doesn't really apply to me anymore. Especially with Spring Term, we always have changing instructors. Like, they might not even teach in the Fall or Winter, or they might even just teach that one time and that's it for them, so sometimes it just doesn't help that much. 2, I also don't really use it that often, unless I have talked to them before about it, and there is still no change I might write that down. (EVF4).

I wouldn't say that I think a lot about it, but I certainly distinguish between the 4 and the 5. Ya, like if I had a really good prof and I would go out of my way to take one of their courses again, then that's a 5. Ya, if it is like an above average prof, like I enjoyed it, but I wouldn't necessarily pursue it, then that would be a 4. Then, obviously, 1 and 2. So I think I actually have a pretty narrow band of what I would consider 3. Like, I think that I am more likely than most people to go to the extreme ends because I want to make my point I guess (MAM3).

Resulting Recommendations

- Recognize that summing scores when you are using categorical-ordinal variables to present averages is problematic (as has been discussed in the literature by Stark & Freishtat, 2014).[8] Instead, some suggest best practice is to provide score distributions as opposed to, or in addition to, averages (See Linse, 2017).

---

[8] For categorical variables we assign a number to a category for the purpose of distinguishing cases from one another. Scores on this type of variable are categorical and only the size of each category may be compared. For example, sex is a categorical variable with two categories (if we are focusing on sex as a binary variable, male & female), there is no inherent more/less or higher/lower to being male or female and we cannot rank, add, subtract, multiply or divide these categories (Schulenberg, 2016).

Part 5: SCP Item Understandings

Guiding Questions

Preamble: In this hand-out, you will find a chart with the proposed questions for the revised course evaluation tool.

We are going to look at the bolded questions in more detail. Let's start with the following item:

'The instructor was a clear communicator.'
1. What does this question mean to you?
2. Explain what factors make an instructor a clear communicator?
3. What factors make an instructor a poor communicator?
4. If you had an instructor who communicated lectures very well, but was very poor with email communication, or face-to-face communication, how do you think you would respond to this question? (i.e., what type of rating would you give them?). Why? Explain.

   a. Would you explain a possibly unclear response (on a rating scale question) in an open-ended comment?

'The instructor returned graded work in a reasonable amount of time.'

1. Explain what you define as a "reasonable amount of time". How do you judge this?

'I knew what I was expected to learn in this course.'

1. What do you think this question is asking/What do you think this question means?
2. What factors do you consider when responding to this question?

Findings

Item 1: The instructor was a clear communicator

Our focus group results suggest there is some degree of consistency in how students interpret this item. Across all six Focus groups, one overarching theme was how well the instructor was able to convey course material so that students could **understand** that material. (It should be noted that the importance of this also emerged when students discussed what factors they consider in evaluating an instructor at the low/high end of the scale, as discussed above). However, two additional interpretations were commonly discussed as well:

Communicates to promote understanding	Perception that clear communicators promote understanding by explaining the material (i.e., course concepts) and making connections for students.	So I think that clear communication is mostly about course material that the professor can convey, especially the relationships between the different concepts in the course material. So, right now I'm in microeconomics class. So the different concepts between

		tax, fiscal policy, government, interrelate to each other, so the prof needs to explain how does the government relate to these different concepts. So, I think that, technically, it's more about the course material and if the prof can convey the ideas really in an unambiguous manner (ARM2)
Communicates expectations	Perception that clear communicators are instructors who convey expectations for graded assignments.	I think about exams, so maybe like they clearly stated what we were responsible for, for the exams. Because there are a lot of times when they go over the slides, they may say a few words about one slide and then spend 10 minutes on another. So if they clearly emphasize that you are responsible for this section, but please make sure you know blah, blah, blah… (AHF6).
Multiple interpretations	Perception that clear communication includes a range of factors: volume, functionality of equipment, accents, availability, whether or not the instructor was engaging, communication of important dates/content changes/reminders etc.).	I would read that and I wouldn't really think about the specific words or ways the instructor says them. I would think of: if they had a thick accent, if their mic is working properly, or if they are speaking really timidly and it is hard to hear, I have had profs that have done all 3 of those things, and it makes it very difficult to hear what they are saying. So, usually, that is what I would think of when I hear if they were a clear communicator (MAM2).

Recommendations

- Most interesting to note, when asked to discuss their reasoning behind a high/low evaluation, students commonly discussed the importance of communication (as evidenced in the 'course delivery' theme). Yet, when students were asked about this specific item, they seemed to have some difficulty summarizing how they define "clear communicator".
- This lack of clarity highlights the importance of ensuring students understand what factors they should consider when responding to this item.
- It might be worthwhile to include additional sub-questions to measure the fundamental aspects of communication that we want to measure (as defined in the literature).

Item 2: The instructor returned graded work in a reasonable amount of time.

When asked about this question, students generally held a common sentiment: they expect to receive graded work back with enough time to make improvements on future class assignments.

Feedback for Improvement	Expectation that feedback should be provided with enough time to process it so that improvements can be made for future assignments/tests.	So, a reasonable amount of time, for me, means, "ok, I need some time to do the next project, so I need to know how I am doing," so that I can communicate with the prof, find out when I can meet with them, and, you know, ask a few questions before the next evaluation is coming. So a couple weeks before the next evaluation is a reasonable amount of time to me. And I don't really care how much time in between there is after I submit the first assignment, but I do expect a couple of weeks in between, before the next assignment is due, so I can improve my performance. (ARF4)
Relative to type of assignment	Recognition that the type of assignment being graded (Multiple choice/Essay) will impact the length of time it takes to receive feedback.	For me, the key word is 'reasonable,' and that leads me to question the quality of the grading. So, for example, I had a midterm where it was all written response and, like, 4 or 3 questions, and we got our marks back literally in two days. So that got me thinking, did they really read everyone's answers word for word, properly? I guess it was just really fast and we were a pretty large group (AHF3).

Implications

- These findings suggest that students' interpretations seem to reinforce that they understand that "reasonable" is a relative term. For students, the term does not necessarily mean a pre-conceived time-frame (24 hours, or one week), but instead means

"enough time to process the feedback so I can do better next time." This interpretation seems to be the intention behind the formulation of the question. Effective teaching gives students the opportunity to practice their learning with feedback.

Item 3: I knew what I was expected to learn in this course

There was some consistency in the language students used to describe their understanding of this item. Interestingly, in their responses students often explicitly used the terminology of "learning objectives" or "teaching outcomes", even though (it is important to note) students were not prompted to use this vocabulary. Despite some consistency, overall, it seems there was some confusion surrounding the interpretation of this question. Students seem to have somewhat different views about what "expected to learn" referred to (e.g., the goals specified by the instructor for the course, the calendar description of the course, the place of the course in the academic program, etc.).

Learning outcomes	Any reference to 'learning outcomes/goals' or 'learning objectives'.	Sometimes they say, like, "learning goals," and put those in their lectures and say, "this is what we are going over today, this is what you should know by the end of the lecture' (SCF1).
Reference to syllabus	Any reference to the syllabus to make sense of this item.	I don't know how relevant this question is for some courses, since some profs may deviate from the syllabus and that may be ok for certain courses. If it is an upper year course, that only this prof teaches, and finds out that the class is more interested in this topic, rather than this topic, and they focus more time on that, then I don't think that is something that is wrong with the course (MAO1).
Reference to assessments	Any reference to graded assignments (i.e., tests, labs, reports etc.)	For example, last term I took chemistry and a chemistry lab, and for part of the chemistry lab the prof said "you don't need to study this for the final," and then the

		questions for the final were from that part (SCF7).
Reference to generic course description (e.g. in the calendar).	Any reference to the generic course description (consulted prior to enrolling in the class).	So, during course selections I would read the little blurb about the course from the Waterloo website, and then I would base off of what I read to the end of the year and what I actually learned (ARF5)
Confusion	Comments that signify confusion. Some students indicated this item lacked clarity.	I actually think this could be taken two separate ways and it is probably not the best wording. I feel like some people would take this as: I knew what I was expected to learn in the course, as in, I already knew a large chunk of material. Or, some people may take it as, I knew what the material was that I was going to learn (MAM1).

Implications/Recommendations

- It is unclear whether students can reliably discern what this question is supposed to measure. Therefore, education or rewording would seem necessary. It does appear that terminology like 'learning objectives' or 'learning outcomes' could be used, but education would be necessary.
- We might want to consider using a software platform with an interactive interface. For example, students could move the cursor to the word 'expected' and a definition would pop up explaining what students should be focussing on with respect to this item (i.e., explain learning outcomes etc.).

<u>Additional Information</u>

Students across all six Faculties indicated they will consult unofficial review sources to learn about instructors.

And to follow up with that last point that was brought up, I understand like the privacy and keeping the evaluations between like the students and the professor, but, like, a different rationale of not completing these surveys are students do not get to have that

information shared between future generations of students. So, I know that students have to default to websites like UWFlow and RateMyProfessor to actually get some of that relevant information 'cause we are not actually receiving that feedback. So, the data we are putting in isn't beneficial at, like, the student level systems, so students have to default to other platforms like RateMyProfessor to kind of like rate teachers (AHF5)

This comment emphasizes the need to consider having some institutional information about SCP surveys made public. It is worth asking whether we would rather have students go to official UW sources to find information about a particular course or instructor, or have them consult external (unofficial, unsystematic) sources to find this information.

The opportunity to fill in open-ended questions is highly valued by students. Students across all six Faculties explained that they use these comments to contextualize their scale selections, particularly when they are at the extreme ends. The following quotes summarize this view:

- If the professor is very good and I have a lot to say in the short text box at the very end, just, like, blah, blah, blah, what they can improve, what they should consider redesigning for the course. And then, on the other side is if there are a lot of issues that I found in the course, then, um, issues that I think the professor should take a look at. And then if they are just average and I do not have much of an opinion, then I would feel more inclined to just go through the, um, the ratings, and less so focus on the short answer (ARF2).
- Definitely, I would write more when I have had a bad experience with a prof compared to a good experience (MAM3).
- If I gave a 1 or a 5, since they are extreme I would explain. Like, I would likely justify myself. Like, if I gave a 5 I would say things that the professor did that were above and beyond, but if I gave a 1 I would say some things they could improve on. (SCF5).

Concluding Remarks

The purpose of the focus groups was threefold: (1) To determine how well students' perceptions of a positive/negative course experience are aligned with evidenced-based principles of effective teaching as defined in the University of Waterloo's Undergraduate Learning White Paper Report and how well these views align with the core questions proposed by CEPT1; (2) To examine how students interpret and understand items from the draft set of core questions and; (3) To understand the response processes of students.

It is important to emphasize that students did not receive a copy of the draft set of core questions until the second half of the focus group session. Thus, students were not primed to define their perceptions of "positive" or "negative" course experiences in terms of the proposed core questions. The findings from this focus group study highlight the following key findings:

- Across all six faculties, there is alignment between student perceptions about

what contributes to a positive course experience and the evidenced-based principles of effective teaching as proposed in the University of Waterloo's Undergraduate Learning White Paper Report and the core questions proposed by CEPT1.

- The following two items, *'The instructor was a clear communicator'* and *'I knew what I was expected to learn in this course'* appear to be somewhat unclear to students and thus should be re-worded to better capture how students understand each item.
- Students use the Likert scale most effectively when all scale points are labelled clearly. Inclusion of a "no response"/ "prefer not to answer"/ "not applicable" option- a comment box to explain this selection may also be valuable.
- Clear descriptions of how to use the scale appropriately should be provided.

The findings from these focus group sessions have been used to inform a number of recommendations, outlined throughout this report. These recommendations will help inform decisions concerning the pilot-test of the core items which will take place during the course evaluation period in the Fall 2018 term.

References

Allan J., Clarke, K. & Jopling, M. (2009). Notions of effective teaching: An exploration of students' and lecturers' perceptions of first year education undergraduate. *International Journal of Teaching and Learning in Higher Education*, 21, 362-372.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine.

Gravestock, P. & Gregor-Greenleaf, E. (2008). *Student course evaluations: Research, models and trends*. Retrieved May, 30, 2018, from the Higher Education Quality Council of Ontario Web site: http://www.heqco.ca/ SiteCollectionDocuments/Student%20 Course%20Evaluations.pdf

Krueger, R.A. & Casey, M.A. (2009). *Focus Groups: A practical guide for applied research*. Thousand Oaks, CA, US: Sage Publications.

Linse, A.R. (2017). Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation* 54, 94-106.

McKeachie, W.J. (1997). Student ratings: The validity of use. *American Psychologist*, *52*(11), 1218-25.

Nasser. F. & Fresko, B. (2002). Faculty view of student evaluation of college teaching. *Assessment & Evaluation in Higher Education*, 27(2), 187-198.

Ory, J.C. & Ryan, K. (2001). How do student ratings measure up the a new validity framework? In M. Theall, P.C. Abrami, & L.A. Mets (Eds.), The student ratings debate: Are they valid? How can we best use them? [Special issue]. *New Directions for Institutional Research*, 109, 27-44.

Theall, M. &Franklin, J. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in students ratings of instruction? In M. Theall, P.C. Abrami, & L.A. Mets (Eds.), The student ratings debate: Are they valid? How can we best use them? [Special issue]. *New Directions for Institutional Research*, 109, 45-56.

Thomas, D.R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation,* 27, 237-246.

Strauss, A. & Corbin, J.M. (1998). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA, US: Sage Publications. Inc.

Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education*, 23, 191-211.

Wagenaar, T.C. (1995). Student evaluation of teaching: Some cautions and suggestions. *Teaching Sociology*, 23(1), 64-68.

Willits, F. & Brennan, M. (2017). Another look at college student's ratings of course quality: data from Penn State student surveys in three settings. *Assessment & Evaluation in Higher Education,* 43(3), 443-462.