Your Course Evaluation Report :
A Guide for Instructors

It is without question that course evaluations are both celebrated and dreaded among instructors. More often than not, myths about course evaluations prevail, with concerns about their reliability, accuracy, meaning, and usefulness, at the forefront of debate (Eisler, 2002). Despite what some believe, the validity and reliability of course evaluations have received much empirical attention. Research confirms that students can be reliable evaluators of various aspects of their course experiences, including the assessment of the presence of teaching behaviours, knowledge gained, difficulty level, workload, and the value of tests, projects, etc (Nasser & Fresko, 2002; Theall & Franklin, 2002).

Assumptions aside, the reality is that teaching, or rather, *teaching well*, is an essential aspect of your academic career at the University of Toronto; and course evaluations are an important part of that assessment.

In the following pages, you will find tools to guide you through the interpretation of your course evaluation report. At times, information may appear repetitive; however, any repetition is simply for emphasis and clarification purposes. You will also note that interpretative tools are provided for each *section* of your report. Bear in mind your report may look different from your colleague's; divisional guidelines specify the content and number of course evaluation items, which can vary across courses and instructors. As you begin your analysis of your course evaluation report, please consider the following general guidelines.

General Guidelines for Reviewing Your Course Evaluation Report

1. Read, think about, and understand the core institutional course evaluation items *before* you review students' responses. As discussed, each core item represents an institutional priority that reflects ideal learning experiences for students in <u>all</u> *University of Toronto courses*.
2. If applicable, read through and understand your divisional and departmental items. These items reflect teaching priorities deemed important for courses in your division and/or department.
3. Review the quantitative data in your report. The following pages outline a step-by-step approach to understanding the statistical summary of students' responses. Statistics in your report not only describe students' responses but provide information regarding their overall validity and reliability.
4. At the onset, note if you have any extreme responses for any item in your report. For example, you might notice a few students felt that the course was "not at all" intellectually stimulating. Whereas it is important to pay attention to all responses, it is also important to note a few extremes are not cause for worry. For example, in a class of 80 students and a response rate of 50 students, one or two comments suggesting the course was "not at all" intellectually stimulating are not reflective of the majority of students' responses.
5. Comparatively, if for an item you notice a majority of students felt the aspect was *not at all* or only *somewhat* part of their learning experiences in the course, this would be an important teaching priority to pay attention to. One of the best ways to respond to this is to be proactive for the next time you teach the course. See Strategies for Enhancing your Teaching located here: http://www.teaching.utoronto.ca/teaching/essentialinformation/evaluation-framework/enhancing-teaching.htm
6. Take time to read through students' responses to the open-ended items (e.g. institutional items 7 and 8). Information provided here might clarify the quantitative data for each item, especially for items you are concerned about.
7. If your course is considered low enrollment or if your course evaluation response rate is between 5 and 9, statistical summary information will not be provided. Rather, response distributions for each item will be provided for your review.
8. Finally, take a step back and notice the extent to which students' quantitative and written information tell a story. Write out a summary for your course. Does this summary correspond with your personal sense of your teaching? If not, why not?

9. **Remember:** If you would like consultation on any aspect of your course evaluation report, please contact Dr. Cherie Werhun at the Centre for Teaching Support and Innovation, cherie.werhun@utoronto.ca.

Understanding the Executive Summary Page of your Report

Section 1 Course Evaluation Executive Summary

Course Name:		Instructor:
Course Code:		

Part A. Core Institutional Items

Mean Responses to Core Items is students' average responses to each core item.

Scale: 1 - Not At All   2 - Somewhat   3 - Moderately   4 - Mostly   5 - A Great Deal

Question	Mean
1. I found the course intellectually stimulating.	4.3
2. The course provided me with a deeper understanding of the subject matter.	4.3
3. The instructor _____ created a course atmosphere that was conducive to my learning.	4.2
4. Course projects, assignments, tests and/or exams improved my understanding of the course material.	3.8
5. Course projects, assignments, tests and/or exams provided opportunity for me to demonstrate an understanding of the course material.	3.6
Institutional Composite Mean	4.1

The **response scale** for items 1-5 requires students to indicate the *extent to which each item was part of their learning experience* in the course. Scores that suggest the core institutional teaching priorities were *mostly* or *a great deal* part of the students' learning experiences in the course are desired. Unless specified, this response scale is used throughout the course evaluation.

The **Institutional Composite Mean** is the average of students' mean responses to the five institutional composite items (i.e., Mean of means for items 1 to 5).

Scale: 1 - Poor   2 - Fair   3 - Good   4 - Very Good   5 - Excellent

Question	Mean
6. Overall, the quality of my learning experience in this course was….	3.9

Note Item 6 is associated with a different response scale that ranges from Poor to Excellent.

Based on ***divisional and unit/departmental guidelines***, the overall summary of your report may have response charts for divisional and departmental items, as well.

Note: Whereas core institutional, divisional, and unit/departmental items are presented for summative purposes, your instructor items are presented for formative purposes only, and are located in the detailed summary pages of your report.
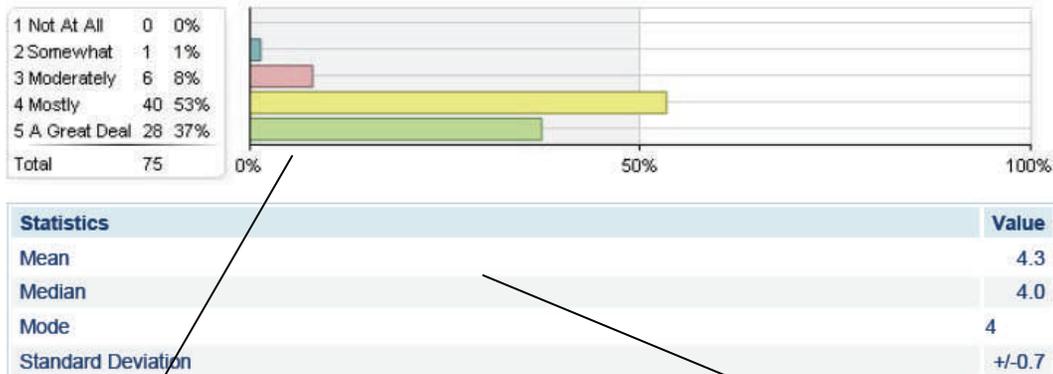
Instructors may opt to share information collected from their instructor-specific items with members of their departments or divisions. Instructors who do not wish to share evaluation data with students must indicate their desire to opt-out.

Section 2 Course Evaluation Details and Summaries
Response Distributions and Associated Statistics

Part A. Institutional Items

1. I found the course intellectually stimulating.

1 Not At All	0	0%
2 Somewhat	1	1%
3 Moderately	6	8%
4 Mostly	40	53%
5 A Great Deal	28	37%
Total	75	

Statistics	Value
Mean	4.3
Median	4.0
Mode	4
Standard Deviation	+/-0.7

Step 1

For each item of the detailed summary page, you will see a bar graph or **response distribution**. Each bar represents the **percentage of student respondents *who endorsed each response of the scale.***

Note the distribution. Is it skewed to one side? In other words, do responses fall to one side of the scale more so than on the other side? Is it very pointy? Do students' responses cluster around one part of the response scale? Or, is it flat or spread out? For some items, students' responses might be highly divergent and you will see "two peaks" in the distribution. Note the percentage of students who selected "not at all" and/or "somewhat". What percentage of students selected "mostly" and/or "a great deal"?
Note this information for each item.

Step 2

The mode, median, and mean provide statistical information about the distribution of student ratings. All three are helpful in the interpretation of your results, especially when the distribution is skewed to one side of the response scale. Note these values for each item.

-The **mode** represents the tallest bar (largest percentage) and is the student response rating that occurs most frequently for that item.
-The **median** represents the middle student response rating when responses are ordered from smallest to largest.
-The **mean** is the average of all students' responses to that item.

Step 3

You'll note for some of the items, the **mean** and the **median** *are not the same.* This is because the mean is influenced by all students' ratings, especially extreme ones (i.e. ratings that are at either end of the response scale) when the sample size is small.

When interpreting the mean, a good question to ask is whether the mean reflects the majority of students' ratings. In other words, does the mean represent the distribution of students' ratings well? One way to assess this is to look at the **standard deviation**.

Generally speaking, a standard deviation around 1 (or less) suggests that a *majority* of student response ratings (approximately 68%) are located within one unit of the mean; and therefore, the mean is a good representation of students' responses. In comparison, large standard deviations suggest that there is a great deal of variability in the ratings, and a majority of responses are dispersed from the mean. In this situation, the mean is unlikely a good representation of student responses and the median is more appropriate.

In sum, a large item mean with a small standard deviation suggests that most students felt the learning aspect was *mostly* or part of their course experience *a great deal*, whereas a small item mean with a small standard deviation suggests that most students felt that the learning aspect was *not at all* part of their course experience. In situations where the distribution of students' responses is highly skewed, the median, or the mid-point of the distribution, is likely a better reflection of students' responses.

The **standard deviation** represents the average *spread* or *distance* of each student's response rating from the mean of response ratings. Small standard deviations suggest that student response ratings are close to the mean whereas large standard deviations suggest ratings are distant from the mean. If all student response ratings were identical, the standard deviation would be zero.