



How professors can use student course ratings to improve teaching and to prepare for tenure/promotion/merit decisions*: Research-based suggestions

Course ratings (also called student evaluations) are often used as an assessment of teaching. Ideally, they form a part of a comprehensive system of faculty evaluation in order to document the full range of teaching responsibilities, c.f., McGill teaching portfolio guidelines. Student ratings cannot stand alone; as data, they provide information about student perceptions, reports of a particular course, that need to be interpreted. The difficulty, as with all data, is how to interpret them in ways that are meaningful and justifiable. The following suggestions are drawn from the literature which represents more than 30 years' investigation into the use and usefulness of student course ratings.

A. Implementation

1. Follow the established departmental procedure. Consider the following.
 - Ensure that students are told what the ratings are to be used for, e.g., personnel decisions, improvement, since students who assume 'personnel decision' will tend to rank higher. Stating this explicitly enables all students to approach the task in the same way.
 - Leave the room during the process as presence tends to increase the ratings; as well, it may make students less inclined to provide written comments.
 - Feel free to distribute the forms at different times in the latter half of the term as this may help students avoid 'rater fatigue', i.e., completing 5 forms in one week. There is no evidence that time of distribution impacts ratings after the half-way point.

B. Interpretation for improvement/personal use

Consider the following when using course evaluations to improve your teaching as well as to present to others your teaching activities and efforts at teaching improvement.

1. Situate the data: In interpreting the results consider descriptive information such as: distribution of responses by item, measures of central tendency (mean, median, standard deviation), direct estimate of error such as confidence intervals for mean.
2. Written comments: These are useful for teaching improvement since they often represent a full range of opinions (usually the most positive and most negative) and may provide insight into why some students had difficulty learning.
3. Nature of course: When considering the meaning of student ratings, do not ignore the nature of the courses since some course factors increase ratings:
 - courses that are elective or chosen from personal motivation
 - courses at higher levels, e.g., 3rd or 4th year, graduate
 - courses smaller than 10 to 15 are ranked most highly, then those 15 to 35 and over 100; least favourably rated are those from 35 to 100
4. Ratings and consultation: Research shows that course ratings are most useful in improving teaching effectiveness when coupled with appropriate consultation. You may want to discuss the meaning of the ratings with a trusted colleague, your Chair or someone from CUTL. If you do so, be sure to include this activity in your teaching portfolio. During the consultation:
 - Focus on specific items, not the global ones. Specific teaching behaviours (e.g., clarity of objectives) are easier to change than personal characteristics (e.g., enthusiasm).
 - Consider ratings in relation to written comments to see if the latter provide indicators/suggestions for improvement.
 - If there is considerable variation in response on an item (e.g., some report assignments as appropriately challenging and others as too challenging.), it may represent important differences in the nature of the students, e.g., senior versus first year.

C. Interpretation for tenure/promotion/merit decisions

1. Remember data provide each student's relative weighting. Their nature makes it imperative to interpret them for others, so that they are not statistically misused.
 - Analysis of results should include descriptive information as noted in Section B, #1.
 - Student ratings can be used to make initial judgments of instructional effectiveness, e.g., exceptional, adequate, unacceptable. They need to be substantiated by other sources of information, e.g., a colleague's assessment of the goals of the course.
 - Given the complexity and variety of courses, disciplines and teaching situations, ratings of global items (e.g., "Overall the professor is a good teacher") only provide general judgments. For a truly comprehensive understanding of a particular course (or your teaching over time), consider the entire pattern of students' responses to all questions.
2. Number of courses: Describe the overall interpretation from a range of courses. Aim for ratings from a variety of courses: research suggests courses from every term for at least two years, totalling at least five courses.
3. Number of respondents: If the number of respondents in a course is small, the ratings should be interpreted with caution. Specifically caution readers if there are fewer than 15 raters in any one course. Ideally, 2/3's of the class should have responded, higher if fewer than 30 students in course.
4. Nature of course: When considering the meaning of student ratings, attend to the nature of the courses since some course factors increase ratings (Section B, #3).
5. Nature of intended learning: If the educational goals of a course or the methods used are not reflected in the questions on the form, note this for readers of your merit submission or teaching portfolio. You may wish to add questions to the form that provide the information you are looking for.
6. Explanation of ratings: Demonstrate how you are using the ratings to improve your teaching.

References

- Cashin, W. (1995). *Student ratings of teaching: The research revisited*. Kansas State University: Center for faculty development and evaluation.
- d'Apollonia, S., & Abrami, P. (1997). Navigating student ratings of instruction, *American Psychologist* (Special issue on student ratings of professors), 52 (11), 1198 - 1208.
- Franklin, J., & Berman, E. (1998). Using written comments in evaluating teaching, *On line Instructional Evaluation and Faculty Development*, 18, 1, Nov.
- Greenwald, A. (1997). Validity concerns and usefulness of student ratings of instruction, *American Psychologist* (Special issue on student ratings of professors), 52 (11), 1182-1186.
- Greenwald, A., & Gillmore, G. (1997). Grading leniency is a removable contaminant of students ratings, *American Psychologist* (Special issue on student ratings of professors), 52 (11), 1209 - 1217.
- Marsh, H., & Roche, L. (1997). Making students' evaluations of teaching effectiveness effective: the critical issues of validity, bias and utility, *American Psychologist* (Special issue on student ratings of professors), 52 (11), 1187-1198.
- McBean, E., & Al-Nassri, S. (1982). Questionnaire design for student measurement of teaching effectiveness. *Higher Education*, 11, 273-288.
- McKeachie, W. (1997). Student ratings: The validity of use, *American Psychologist* (Special issue on student ratings of professors), 52 (11), 1218-1225.
- Theall, M. (1998). Colloquy: An Editorial, *On line Instructional Evaluation and Faculty Development*, 18, 1, Nov.
- Theall, M., & Franklin, J. (1991). Using student ratings for teaching improvement, In M. Theall & J. Franklin (Eds.) *Effective practices for improving teaching*. San Francisco: Jossey Bass, 83-96.

*Companion document available from CUTL: *Using student course ratings for personnel decisions*.

**CUTL Website – <http://www.mcgill.ca/cutl>