



RED DEER COLLEGE
Faculty Performance Committee

Interpreting Student Feedback
Instrument (SFI) Results
April 2019

Interpreting Student Feedback Instrument (SFI) Results

Faculty Performance Committee

April 17, 2019

Red Deer College's Faculty Performance Policy mandates the use of a standardized student feedback instrument. The Faculty Performance Committee (FPC), comprised of both faculty and administrative representatives, recognizes that the use of student feedback instruments (SFIs) can be contentious and stressful for faculty members. The FPC has conducted a review of SFIs, which included looking at practices at other post-secondary institutions, consulting the relevant literature, reflecting on RDC's past practices with student feedback instruments, and gathering input from constituency groups.

As a result of this process, the FPC remains committed to having a standardized student feedback instrument as a tool for assessing and improving instructional effectiveness of RDC's faculty, and have extensively revised the SFI instrument and how the feedback results will be reported. The FPC has also drafted this document to assist all users of SFI feedback (faculty, deans and associate deans, and members of continuous appointment committees) to promote consistent, balanced interpretation and use of SFI feedback.

Preamble

In accordance with the *Faculty Performance Policy*, all faculty members must engage in activities to assess performance. Student feedback instruments (SFIs) are designed to give students the opportunity to provide feedback on the instruction of a course. In turn, that feedback can be used to improve course delivery and enhance an instructor's teaching practice. Although SFIs (as with any form of feedback reporting tool) have limitations, RDC places great, but not exclusive value on the feedback the SFI provides. SFIs are only one part of a faculty member's performance package, and should be used with other forms of assessment when evaluating faculty performance.

Interpreting numerical data

Student Feedback Instruments are used to gather the *collective* views of students about their experience in a course. Therefore, no individual measure such as a score, a mean, or a percentage should be overemphasized. Rather, the whole distribution across the instrument should be the source of information. As such, the focus of analysis should be primarily on any *trends* evident in the numerical results. Any trends in numerical results can also be useful for comparison from year to year or from one course or section to another.

Other factors to consider when analyzing results include overall response rate, and the context or characteristics of the course. For instance, a literature course may be expected to have a lower rating on a 'real world application' question, whereas a nursing course may have higher rating in this area.

Interpreting the mean

Despite its common usage and broad appeal, the mean that is calculated from the SFI responses must be interpreted with caution. For instance, *strong agreement* can mean different things to different people. Therefore, the mean of all scores on any one item does not have the objective quality that is often imputed to it. Moreover, the mean by itself may hide some important aspects that only the distribution can reveal.

For instance, the following fictitious SFI results generate a mean of 3.7, as explained here:

	Strongly Disagree				Strongly Agree	Totals
Score	1	2	3	4	5	
# of Respondents	1	2	4	8	5	20
Score×# of Respondents	1	4	12	32	25	74
	Mean = 74/20 = 3.7					

However, the following results would produce the same mean, while indicating a rather different class response:

	Strongly Disagree				Strongly Agree	Totals
Score	1	2	3	4	5	
# of Respondents	3	3	0	5	9	20
Score×# of Respondents	3	6	0	20	45	74
	Mean = 74/20 = 3.7					

Finally, the value of the mean may be affected by many factors that may not reflect the instructor’s performance.

Therefore, the mean should be used only as an imperfect summary and as an indicator, rather than a final assessment of teaching effectiveness. While unusually high or low means should lead to further analysis, be it in comparison to institution-wide summaries or to previous terms, they should not be considered as sole criterion measurement on which an instructor’s performance is assessed.

This advice is informed by the 2018 arbitration decision from Ryerson University which concluded that mean scores from student feedback instruments are so flawed as measures of teaching effectiveness that they should not be used for this purpose. The arbitrator noted that teaching dossiers, peer observations, student comments, frequency distributions and response rates, should all be employed in assessing teaching effectiveness, particularly for promotion and tenure decisions.ⁱ

Interpreting student written comments

As such, numerical responses should not be used as the sole means of analyzing SFI results. Student comments can provide richer feedback, and help deepen the understanding and interpretation of numerical responses. During analysis, the focus should remain on identifying common or emergent themes within the student comments. Comments can be useful in supporting or further clarifying numerical results, as well as identifying specific teaching practices that may enhance the learners’ progress.

Experience suggests that almost all faculty members at some point receive negative feedback. The presence of such feedback should not be overvalued, but its content, validity and frequency should be properly considered. Undue criticisms or personal attacks are not constructive, and should be given little consideration when reflecting on an instructor’s performance.

Interpreting Atypical Results

It is important not to give undue weight to unusually high or low results in a given class. Every faculty member, even a highly skilled one, may receive an occasional low average rating. Inconsistent SFI results may stem from factors such as trying new teaching approaches, or unusual characteristics or attributes of a class. Such circumstances warrant careful review and reflection, but in considering the numerical results, focus should be placed on the patterns of scores over time.

Common Concerns

Is student feedback valid and reliable?

The validity of an instrument is its ability to measure what is supposed to be measured, while the reliability of an instrument is its ability to measure consistently. As some aspects of teaching and learning are inherently easier to measure than others via questionnaire, and student responses are affected by many temporary and external factors, SFIs are not completely reliable or valid. However, research indicates that student evaluations “are a more reliable and valid representation of teaching quality than any other method” (b, p. 97).

SFIs give students an opportunity to communicate and reflect on their experience in a class (*b*, p. 96). There are known concerns with students assessing particular elements of an instructor's performance, such as the instructor's expertise within the discipline (*a*, p. 33). However, aside from the instructor, the students are the only ones to have experienced the course as a whole, and therefore are well-positioned to comment on facets such as organization, alignment of content and assessment, and effective learner engagement (*a*, p. 33).

If student feedback is controversial and may be subject to bias, why do we keep using it?

There are few alternatives that provide such "quantifiable and comparable" data (*a*, p. 9), and research indicates that student evaluations are a useful measure of teaching effectiveness (*a*, p. 9; *b*, p. 94).

Any qualitative feedback based on human observations is vulnerable to bias by the observer or the reviewer. It is almost impossible to draw conclusions of whether results are biased, and even more challenging to identify specific reasons for any bias that may be present (*b*, p. 98). Overall, research shows either little significant correlation or inconclusive results with regards to biasing factors (*a*, p. 47-50).

As stated previously, it is strongly recommended that SFI evaluations not be used in isolation (*b*, p. 100). As described in the *Faculty Performance Policy*, SFIs are one of several measures used to inform and improve teaching.

What influences can SFIs have on subsequent instruction?

Evidence suggests that changes in the classroom are more likely to happen when faculty feel supported and motivated (*a*, p. 16). Faculty members are therefore encouraged to seek support and discuss their results with colleagues, mentors, or with CTL faculty.

Is "easy" grading correlated with positive feedback?

In the research, there is a weak relationship between higher grades and more positive student feedback, but it is difficult to make a direct link between just these two factors. The existence of a significant link is not generally accepted by researchers (*b*, p. 97).

Moving Forward

After reviewing SFI results, take some time to reflect and look for patterns and trends in students' comments. Note what has been done well and celebrate the positive. Acknowledge that some feedback can be unhelpful or even hurtful, but that constructive feedback can provide guidance for continued growth. For any actionable suggestions for improvement, consider specific, goal-based adjustments that can be made. The faculty member is encouraged to discuss the feedback with others in order to better achieve understanding, as well as to brainstorm ideas and share best practices.

The Faculty Performance Committee believes that by following the guidelines as detailed in this document, student feedback will continue to support and nurture excellence in teaching at Red Deer College.

Sources:

^aGravestock, P. & Gregor-Greenleaf, E. (2008). *Student course evaluations: Research, models and trends*. Toronto, Ontario: Higher Education Quality Council of Ontario. Retrieved from: http://www.heqco.ca/SiteCollectionDocuments/Student%20Course%20Evaluations_Research,%20Models%20and%20Trends.pdf

This large-scale literature review looked at hundreds of documents, and at its publication in 2008, was the first synthesis of the "research on student course evaluations from a Canadian perspective" (p. 7). The authors (both affiliated with the University of Toronto) note the divisive nature of the scholarship on student evaluations and the bias evident even amongst scholars, although some facets around evaluations are supported strongly by research. The authors also note

the difficulty inherent in research on evaluations, given the enormous variety of evaluation instruments, teaching and administrative practices, and institutional contexts. Nonetheless, the authors make the best attempt possible to synthesize the research and “incorporate all the major themes we identified in the research” (p. 8).

^b 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. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0191491X16300232#sec0015>

Although drawing upon a vast body of work related to student evaluations, the aim of this author is not to offer a literature review, but to instead dispel some of persistent myths around student evaluations that are not supported by the literature. The author’s target audiences are those in administrative roles who use/mandate the data collected about faculty members. The author argues that the key research-based findings are not reaching faculty and administrators, who rely instead on mainstream and/or academic sources that perpetuate misinformation. The article outlines a number of suggestions on how to best and fairly interpret/use data from student evaluations, and how to counter the negativity around student evaluations using research-based practice.

ⁱ Ryerson University v Ryerson Faculty Association, 2018 CanLII 58446 (ON LA), <<http://canlii.ca/t/hsgkz>>, retrieved on 2019-04-01