

focus

on university teaching and learning

Volume 6, Number 5

May/June 1997

Student Ratings at Dalhousie

At the end of the 1997 winter term, faculty, students, and staff from the Faculty of Science and the School of Business Administration took part in the Student Ratings of Instruction Pilot Project which tested a proposed standardized system for gathering feedback from students about their perceptions of the quality of teaching in classroom and laboratory settings at Dalhousie University. This issue of FOCUS describes the project, its aims, and its outcomes.

Using Student Ratings of Instruction

Student ratings of instruction (sometimes called "course evaluations" or "instructor evaluations") have long been used for a number of purposes: as evidence of teaching performance for personnel decision-making, as information for the instructor's use in improving teaching and learning, as an aid to students in selecting courses, and as data used for purposes of accountability and the maintenance of standards

Student ratings practice has been heavily researched, with over 70 years of scholarly work on the subject (University of Massachusetts, 1995). As a result, there is clear evidence that student ratings of instruction, *properly collected and interpreted*, provide valid, reliable information on teaching performance. However, not all types of student ratings are suitable for all purposes. To ensure fairness in administrative processes, a fundamental distinction must be made between summative

(for personnel decisions and other institutional assessments) and formative (for teaching improvement) evaluation.

Summative versus Formative Evaluation

Summative evaluations of teaching are concerned with the overall quality of teaching performance which is assessed as required for personnel management or for purposes of documenting the quality of institutional activities. Summative student ratings of instruction focus on those teacher behaviours that are correlated with desirable outcomes for students and with other measures of effective teaching. Because of the high correlation between this "summary" or "global" evaluative material and measures of student learning, such data is the most appropriate for personnel decision-making or making judgements about quality.

Formative evaluation is a flexible process usually initiated and controlled by the professor with a view to making changes in teaching practices in order to enhance student learning. Formative student ratings of instruction focus on the minutiae of teacher behaviour, course format, and curriculum content. Because this is a diagnostic process undertaken for developmental purposes, formative evaluation should be done regularly throughout the course, so that changes can be made which will benefit the students in the class. Standardized, multi-dimension student ratings forms can be a useful way to gather such information, but because the results obtained have a low correlation with measures of student learning, their use in summative evaluations is inappropriate (Braskamp et al., 1984; Abrami and

d'Applonia, 1990; Arreola and Aleamoni, 1990).

A review of current practices at Dalhousie suggests that the distinction between summative and formative evaluation is often blurred. Many departments employ dual-purpose questionnaires which clearly differentiate between information collected for summative purposes and information which will be used by the professor to improve the course. However, a number of other departmental questionnaires are a confused mixture of summative and formative items. A separate end-of-term summative student ratings process and a variety of teacher-initiated methods of obtaining formative student feedback at different times in the course is the preferred practice.

Student Ratings of Instruction Pilot Project

Purpose and terms of participation

The Pilot Project was initiated by Tom Traves, President, and Deborah Hobson, Vice-President (Academic & Research), and was coordinated by the Office of Instructional Development and Technology. Dean Warwick Kimmins agreed to test the proposed system in the Faculty Science. Five classes from the School of Business also participated.

In a memo to members of the Faculty of Science, Drs. Traves and Kimmins explained the underlying rationale of the Project:

"It is our expectation that by improving the ways in which we evaluate teaching at Dalhousie, we can demonstrate our commitment to excellence as an educational institution. In addition, by improving the quality of evidence used to evaluate pedagogical achievements, we can ensure that faculty are recognized and rewarded for their contributions to our teaching mission. While student ratings of instruction should never be used as the sole indicator of teaching quality, they are nonetheless a source of valid, reliable evidence and every effort should be made to ensure that the data so obtained is properly collected and utilized."

Currently at Dalhousie, the collection, interpretation, and use of student ratings data are

undertaken using a variety of formal and informal mechanisms and policies. Such variation presents problems in terms of ensuring the quality of practice and of the information collected—quality essential in summative or overall performance evaluations connected to personnel decision-making. For this reason, the Student Ratings of Instruction Pilot Project was initiated to test a proposed standardized system for gathering feedback from students on teaching quality. The Pilot Project relates **only to student ratings of instruction used for summative purposes.**

At Dalhousie, administrative use of student ratings data used to evaluate the teaching performance of a member of faculty is governed by a clause in the Collective Agreement between the Board of Governors and the Dalhousie Faculty Association which stipulates that student ratings "gathered as a result of any collective teacher evaluation authorized by the Senate or any Faculty of Dalhousie University shall not be considered anonymous material . . ." which is barred from personal files (Article 18.09, p.43). Because the questionnaire used in the pilot project was not "authorized by the Senate or any Faculty," participation was voluntary and there were strict limitations on the use of the results. In a memo to members of the Faculty of Science, President Tom Traves and Dean Warwick Kimmins clarified these limits:

"Because this is a test of the [proposed] system, it is not our intention that the results obtained in the pilot phase be used in any personnel decision-making process. However, so as not to penalize those applying for tenure or promotion, a faculty member may voluntarily submit the results from his or her own classes as evidence of teaching performance. Further, no administrative body of the university will introduce or use the results from the pilot project to assess the performance of any individual or in any personnel decision-making or record-keeping process, except at the written request of the faculty member concerned."

Distribution of results from the Pilot Project was therefore carefully controlled. A single report was produced and given to each participating teacher for each of his or her classes. No one else received these results nor will they be given to anyone without the teacher's written permission. Each department head received a report on the aggregate results for the department (not for individuals) and copies of these departmental reports were given to the Dean of Science, the Vice-President (Academic and Research), and the President.

Participants

The Pilot Project included 153 undergraduate classes in the Faculty of Science and 5 in the School of Business for a total of 190 instructor/class units (the larger number reflects classes taught by more than one person).

For methodological reasons, some classes were intentionally excluded from the test. In general terms, the reliability of results depends on the number of respondents—the more students, the higher the reliability. For this reason, only classes with enrollments of 10 or more were asked to participate. As well, classes with more than four instructors were not included.

Methods

The pilot project, which was coordinated by the Office of Instructional Development and Technology, incorporated the following principles:

- The pilot project should be concerned with testing a standardized system for gathering student ratings for summative purposes (providing evidence for personnel management and for "accountability" and to aid in planning).
- The system should use a questionnaire which would be appropriate for a variety of teaching situations including classrooms, laboratories, and clinical settings.
- The procedures for collecting data should be based on the findings of research on student ratings of instruction.
- The questionnaire should contain items on those dimensions of teaching which are correlated with positive student outcomes and which are within the ability of students to fairly judge.
- The questionnaire should contain statements about the purposes the ratings are meant to serve and about the procedures for distribution of results.
- The report on results should contain comparative data from an appropriate reference group.
- The report on results should be easily understood and should contain information on response rates, frequency distribution, means, and information on students' motivation for taking the class.

(Cont'd on page 7)



Student Rating of Instruction Questionnaire

DEPARTMENT: _____
 CLASS: _____ SECTION: _____ YEAR: _____ TERM: _____
 INSTRUCTOR: _____

Sample

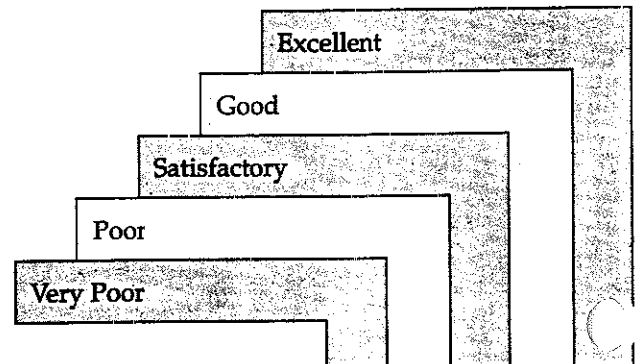
IMPORTANT DIRECTIONS FOR MARKING ANSWERS

Use an HB or soft lead pencil
 Do NOT use ink or ballpoint pens.
 Make heavy black marks that fill the circle completely.
 Erase clearly any answer you wish to change.
 Make no stray marks on the answer sheet.

CORRECT ● INCORRECT ⊗ ⊙ ⊖ ⊕

The information you give on this form will be used to review the effectiveness of your instructor's teaching. Your thoughtful ratings on the questions will be used in the faculty tenure and promotion decision-making process, for other personnel decisions, and to provide ongoing information on teaching effectiveness at Dalhousie University. Your response is anonymous. Your instructor, the Chairperson, and other relevant personnel committees will be provided with a summary of the class responses but will not see any of the completed questionnaires. Your response will not affect your grade. The summary of responses will be given to the instructor only after the final grades have been submitted.

Rate your instructor's performance on each of the following items by marking the appropriate bubble to the right of each item. If you are unable to make a judgement about a particular item, leave that section blank. If you wish to comment on any aspect of the course, please do so on the sheet provided.



STIMULATION OF LEARNING

1. The instructor presented the class in such a way that the subject matter became intellectually stimulating and interesting for you.

VP	P	S	G	E
----	---	---	---	---

ORGANIZATION

2. The instructor organized the class material and presented the individual classes well.

VP	P	S	G	E
----	---	---	---	---

COMMUNICATION

3. The instructor communicated clearly with students.

VP	P	S	G	E
----	---	---	---	---

ENTHUSIASM

4. The instructor showed interest and enthusiasm for teaching the subject matter of the class.

VP	P	S	G	E
----	---	---	---	---

FAIRNESS

5. The instructor was fair and reasonable in evaluating and marking student work.

VP	P	S	G	E
----	---	---	---	---

FEEDBACK

6. Students were given meaningful and timely feedback on assignments and tests (considering any limitations due to class size).

VP	P	S	G	E
----	---	---	---	---

CONCERN FOR STUDENTS

7. The instructor showed genuine concern for students.

VP	P	S	G	E
----	---	---	---	---

OVERALL TEACHING EFFECTIVENESS

8. Compared with other university instructors you have had, how would you rate the instructor's overall teaching effectiveness?

VP	P	S	G	E
----	---	---	---	---

REQUIRED/ELECTIVE

9. In my program, this class was: A. required B. elective

A	B			
---	---	--	--	--

MOTIVATION FOR TAKING THIS CLASS

10. Please indicate the primary reason you took this class. (Choose only one: You may have a number of reasons, but please select only the most important.)

- A. I took this class only because it was required.
- B. I was interested in the subject matter.
- C. I wanted to take a course from this professor.
- D. I took this class because my first choice was full or didn't fit my timetable.

A	B	C	D	
---	---	---	---	--

DALHOUSIE UNIVERSITY STUDENT RATINGS OF INSTRUCTION REPORT

Instructor:

Department:

Course:

Section: 01 Year: 1996 Session: Winter Term: B

Total number of enrollments: 143

Total number of returned evaluations: 78 Response Rate: 54.55%

Number of responding students for whom this class is required: 20 [25.64%] elective: 54 [69.23%]

Motivation of students for taking this class

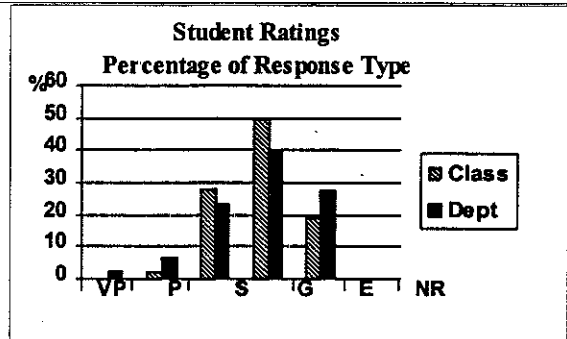
Percentage

I took this class only because it was required:	5.13%
I was interested in the subject matter:	79.49%
I wanted to take a course from this professor:	2.56%
I took this class because my first choice was full or didn't fit my timetable:	11.54%

Notes: VP = Very Poor (1), P = Poor (2), S = Satisfactory (3), G = Good (4), E = Excellent (5), NR = No Response

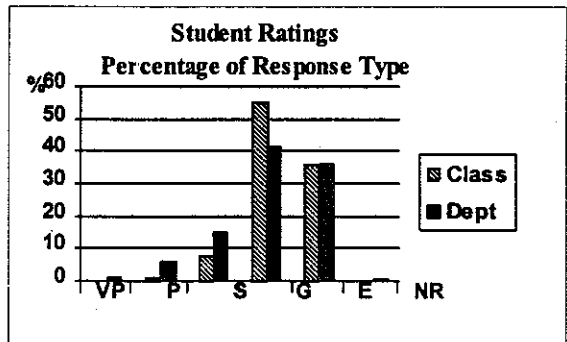
1 The instructor presented the class in such a way that the subject matter became intellectually stimulating and interesting for you.

Mean for this class:	3.9												
Department mean:	3.8												
Lowest mean for a class in this department:	2.2												
Highest mean for a class in this department:	5.0												
Number of respondents	<table border="0" style="display: inline-table; vertical-align: middle;"> <tr> <td>VP</td><td>P</td><td>S</td><td>G</td><td>E</td><td>NR</td> </tr> <tr> <td>0</td><td>2</td><td>22</td><td>39</td><td>15</td><td>0</td> </tr> </table>	VP	P	S	G	E	NR	0	2	22	39	15	0
VP	P	S	G	E	NR								
0	2	22	39	15	0								



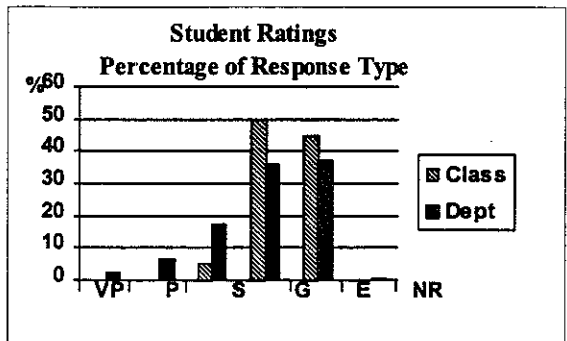
2 The instructor organized the class material and presented the individual classes well.

Mean for this class:	4.3												
Department mean:	4.1												
Lowest mean for a class in this department:	2.6												
Highest mean for a class in this department:	4.8												
Number of respondents	<table border="0" style="display: inline-table; vertical-align: middle;"> <tr> <td>VP</td><td>P</td><td>S</td><td>G</td><td>E</td><td>NR</td> </tr> <tr> <td>0</td><td>1</td><td>6</td><td>43</td><td>28</td><td>0</td> </tr> </table>	VP	P	S	G	E	NR	0	1	6	43	28	0
VP	P	S	G	E	NR								
0	1	6	43	28	0								



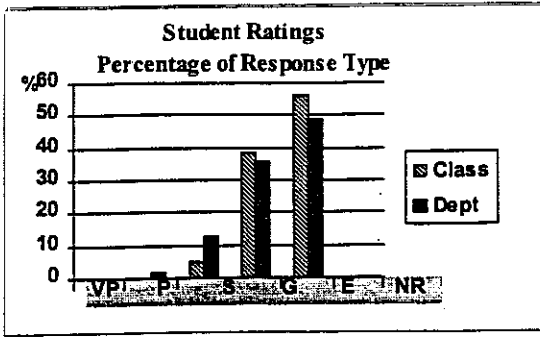
3 The instructor communicated clearly with students.

Mean for this class:	4.4												
Department mean:	4.0												
Lowest mean for a class in this department:	2.0												
Highest mean for a class in this department:	5.0												
Number of respondents	<table border="0" style="display: inline-table; vertical-align: middle;"> <tr> <td>VP</td><td>P</td><td>S</td><td>G</td><td>E</td><td>NR</td> </tr> <tr> <td>0</td><td>0</td><td>4</td><td>39</td><td>35</td><td>0</td> </tr> </table>	VP	P	S	G	E	NR	0	0	4	39	35	0
VP	P	S	G	E	NR								
0	0	4	39	35	0								



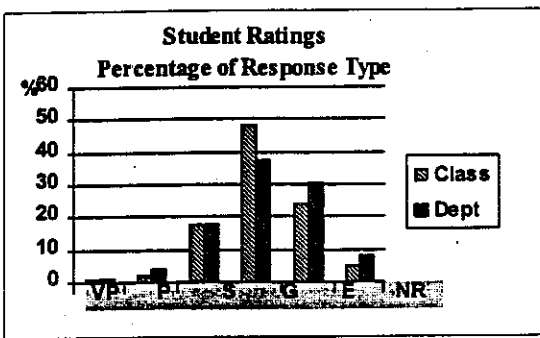
4 The instructor showed interest and enthusiasm for teaching the subject matter of the class.

Mean for this class:	4.5					
Department mean:	4.3					
Lowest mean for a class in this department:						3.4
Highest mean for a class in this department:						5.0
Number of respondents	VP	P	S	G	E	NR
	0	0	4	30	44	0



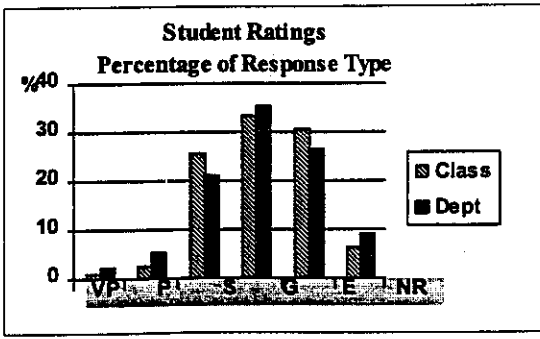
5 The instructor was fair and reasonable in evaluating and marking student work.

Mean for this class:	4.0					
Department mean:	4.0					
Lowest mean for a class in this department:						2.9
Highest mean for a class in this department:						5.0
Number of respondents	VP	P	S	G	E	NR
	1	2	14	38	19	4



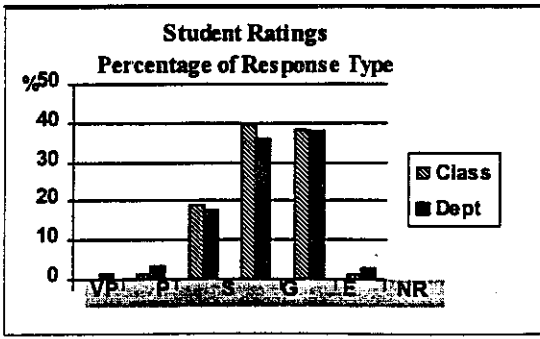
6 Students were given meaningful and timely feedback on assignments and tests (considering any limitations due to class size).

Mean for this class:	4.0					
Department mean:	3.9					
Lowest mean for a class in this department:						2.7
Highest mean for a class in this department:						5.0
Number of respondents	VP	P	S	G	E	NR
	1	2	20	26	24	5



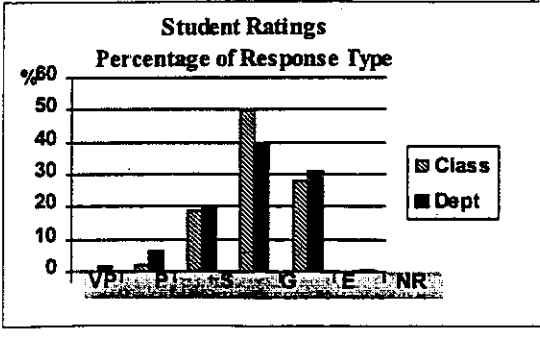
7 The instructor showed genuine concern for students.

Mean for this class:	4.2					
Department mean:	4.1					
Lowest mean for a class in this department:						2.7
Highest mean for a class in this department:						4.8
Number of respondents	VP	P	S	G	E	NR
	0	1	15	31	30	1



8 Compared with other university instructors you have had, how would you rate the instructor's overall teaching effectiveness?

Mean for this class:	4.0					
Department mean:	3.9					
Lowest mean for a class in this department:						2.3
Highest mean for a class in this department:						4.8
Number of respondents	VP	P	S	G	E	NR
	0	2	15	39	22	0



The Student Rating of Instruction Questionnaire consisted of two parts: the first a machine-read form with 10 questions (see p 4) and the second a sheet for written comments. For each participating class (190 in the Faculty of Science and 5 in the School of Business) an envelope containing sufficient questionnaires and directions for their administration was distributed through the departments. Essential elements of the procedures were:

- The teacher appointed a student to oversee the distribution and collection of questionnaires, read instructions to the students, and then left the room.
- The student assistants distributed the forms and put completed and uncompleted forms back in the envelope, which they then sealed and signed. They recorded the number of completed questionnaires and returned the envelope to the department office.

The completed forms were scanned and the data analysed using software developed by staff at Academic Computing Services in consultation with the Office of Instructional Development and Technology. Reports of the results were distributed in the manner explained above. For your information, a copy of one class report is reproduced here with the permission of the instructor (see pp. 5-6).

The class report is designed to provide different ways of looking at the student ratings data. For each question, the mean response and frequency distribution for the class is recorded. Comparative data is provided in three ways: 1) the departmental mean (the average of the means of all participating classes in the department); 2) lowest and highest means for a class in the department; and 3) a bar graph which shows the percentage of response type for the class and the department. Also recorded are information on the response rate, percentages of respondents for whom the class was required or elective, and the primary motivations of students for taking the class.

Dealing with students' written comments has been somewhat more complicated. Ideally, the instructor should receive only a typed version of the comments so that students will feel free to answer without fear of being identified through their handwriting. For the Pilot Project, comments were to be typed by department support staff but limited resources

meant this was not always possible. One department opted out of typing at the beginning, informing students at the time that the instructor would receive photocopies of written comments with any signatures masked. In other departments, typing comments is already the practice, so the pilot represented no change.

Observations

Throughout the pilot project, participants were encouraged to provide feedback to the project coordinator and to suggest improvements. Overall, the response has been positive and many excellent suggestions for improvement have been received.

Experience with the distribution and analysis of the questionnaires proved that a campus-wide system could be efficiently run. For the most part, participants carefully followed procedures designed to ensure the confidentiality and integrity of the data. Some enterprising faculty members put the "Instructions to Students" on an overhead transparency, rather than simply read them aloud as requested. Student assistants did a good job in keeping track of the completed and uncompleted questionnaires and in returning the signed and sealed envelopes. The specially designed software performed very well and the reports generated have been well received by faculty members.

We have discussed the project with a small number of students who have responded favourably. They especially like the length of the questionnaire, the clear instructions, the procedures used. They also like the fact that the way in which the ratings would be used was clear—information not always provided in existing practice. While impossible to confirm, support staff in two departments suggest that it may be the case that a greater percentage of students took the opportunity to make written comments than had been the case with the departmental system. It may also be that the number of signed comments is increased.

The primary aim of the proposed standardized system is to improve the quality of the data used for summative purposes. Other benefits can also be realized. A standardized approach with centralized administration eliminates duplication of effort and frees

departmental resources. It provides a larger information base which can be used to generate better comparative data or to investigate questions of reliability and validity. The aggregate data can be used to demonstrate the overall quality of teaching in a department or faculty. Standardized practices which take the student perspective into account can also increase students' confidence in the ratings system—sending the message that their input is important and that the university is concerned about the quality of their learning experience.

The Pilot Project demonstrated the feasibility of a standardized student ratings of instruction system. Ongoing discussions will determine future practice. We welcome your comments and suggestions.

To comment on the project or the proposed student ratings of instruction system, please contact:

Carol O'Neil,
Research and Programme
Coordinator
Office of Instructional Development
& Technology
(902) 494-1894 or
carol.o'neil@dal.ca

References:

- Abrami, P.C. and d'Applonia, S. 1990. The Dimensionality of Ratings and Their Use in Personnel Decisions. In M. Theall and J. Franklin, Eds., *Student Ratings of Instruction: Issues for Improving Practice*. San Francisco: Jossey Bass.
- Arreola, R.A. and Aleamoni, L.M. 1990. Practical Decisions in Developing and Operating a Faculty Evaluation System. In M. Theall and J. Franklin, Eds., *Student Ratings of Instruction: Issues for Improving Practice*. San Francisco: Jossey Bass.
- Braskamp, L.A., Brandenburg, D.C. and Ory, J.C. 1984. *Evaluating Teaching Effectiveness: A Practical Guide*. Beverly Hills: Sage.
- University of Massachusetts. 1995, October. *Assessment Bulletin*, 1 (2)
-

focus

on university teaching and learning

Focus is the bulletin of the Office of Instructional Development and Technology at Dalhousie University.



M. Carol O'Neil,
Associate Editor
Alan Wright, Ph.D., Editor
Production: Dalhousie Graphics
Office of Instructional
Development & Technology
Dalhousie University
Halifax, Nova Scotia, Canada
B3H 3J5
Tel (902) 494-1622
Fax (902) 494-2063
E-Mail OI.DT@is.dal.ca