

THE USE OF TEACHING EVALUATIONS BY PROFESSORIALATE IN BUSINESS SCHOOLS IN ATLANTIC CANADA

Judy Ann Roy, PhD
Faculty of Business Administration
University of New Brunswick

Student opinion surveys remain the primary tool used to assess the quality of teaching in higher education. This study reveals the relationship between two fundamental needs related to the assessment of teaching behaviours: the institution's need to assess the quality of individual teaching and the instructor's need for feedback.

Background

Educational quality and accountability are two essential characteristics currently used to assess postsecondary institutions. At the institutional level, universities must account for: the standards used to measure learner success and failure rates; outputs and outcomes assessment; comparable management statistics; efficiency and effectiveness of program offerings; a rationalization of programs; compliance with international standards and norms; and operation of facilities and infrastructure provided by public and private funds. Outputs and outcomes assessment includes using student assessments as a tool to measure and evaluate the teaching output of a higher learning institution (CMEC, 1999). At the individual level, feedback from the learner to the instructor should have an operational significance within the institution, providing effective and timely feedback on the quality of learning opportunities and outcomes. To help ensure satisfaction, individual faculty interaction with students and feedback from students are necessary to optimize the educational outcomes (CMEC, 1999).

Institutions and faculty look at quality from many dimensions: quality of offerings, quality of inquiry and research, quality of student life, quality of institutional management, quality of learner outcomes, and quality of instructional delivery. This study focused on features of the teaching evaluation process within the boundaries of the participating institutions and faculties / departments / schools of business. The study assumed that the quality of instruction is improved when the instruction is adequately evaluated and instructors are provided with feedback that will improve their teaching.

The evaluation of university teaching serves two functions: the feedback from the evaluations is used *summatively* – to assess faculty performance for tenure and promotion decisions; or *formatively* – to improve teaching (Centra, 1993; Murray, 1997a). The most widely used method to evaluate university teaching is the student opinion survey (Seldin, 1999) however, often the information captured on student rating forms does not provide concrete suggestions as to how to make changes to improve the quality of teaching. Following is a literature review which looks at four related bodies of literature: the history of evaluating teaching in higher education; the nature and use of student evaluation of teaching; faculty evaluation programs; and effective teaching practices in higher education.

History of Evaluating Teaching in Higher Education

Teaching effectiveness is evaluated a number of different ways including: student evaluations, ratings from former students, instructor self-evaluations, and peer assessments. Of these methods, the most prominent is student evaluations (Seldin, 1999). Student evaluation of teaching has been extensively researched since the early 1970s; however, it has an important history dating back to the pioneering research conducted by Herman H. Remmers in the 1920s.

Four distinct periods of student evaluation research are evident with the first period, 1920-1960 dominated by Remmers and his colleagues. They promoted the use of student ratings, although such data were rarely used until the 1960s. During the second period from 1960-70, student rating forms were used by faculty, however, almost entirely on a voluntary basis. Faculty members would administer the student evaluation forms themselves so that they might make improvements on their own. The third period, 1970-80, has been called the “golden age of research on student evaluation” (Centra, 1993, p.50). During this period a new wave of research studies investigated issues of potential bias, validity and utility, and supported the use of student ratings for formative and summative purposes. Marsh and Dunkin (1997) point out that many published studies in the 1970s were methodologically unsound, but the quality of research and published articles improved in the 1980s. The fourth period, from the mid-1980s to the present, is characterized by the continued refinement and clarification of research results with a series of reviews and meta-analyses that substantiate earlier findings (Centra, 1993).

Since the 1990s, the concept of the scholarship of teaching has taken hold (Boyer, 1990). Boyer’s pivotal work was expanded by Glassick, Huber and Maeroff (1997) who note that the idea of teaching as a scholarly activity has led faculty members to view their teaching work in a more professional way. Research in the areas of scholarly teaching and the scholarship of teaching and learning is the most recent trend in this expanding body of literature (Isaacson, 2000; Knapper & Cranton, 2001, Smith 2001a, Theall & Centra, 2001).

Student Evaluation of Teaching

During the 1990s, the use of student evaluation has increased and has become the preferred source for information on the quality of teaching (Seldin, 1999). Researchers caution against the use of data from only one source when drawing conclusions about teaching quality. Five major concerns related to student evaluation research are: the multidimensionality of ratings, the reliability of ratings, the validity of ratings, potential sources of bias, and the utility of student ratings. Student rating forms are considered multidimensional as they attempt to measure different aspects of effective teaching which is itself a multidimensional construct. Even though no single, agreed-upon list of teaching qualities exists, general characteristics of good teaching, such as strong organizational and communication skills, are reasonably consistent. Current research shows that, when teacher evaluation forms are properly constructed and administered, the ratings, if interpreted with appropriate caution, are undeniably helpful in identifying exemplary teachers and teaching (Feldman, 1997) as well as strengths and weaknesses in teaching behaviors.

Research shows that as the number of raters increase, so does the reliability of the ratings. Reliability refers to the consistency among raters on specific items on the student evaluation form and the stability in the ratings of the same instructor across different courses and over time. Validity refers to the integrity and appropriateness of the conclusions drawn from the evidence.

The validity of student ratings is supported by comparing them to other sources of evaluative data on an individual's teaching, such as self-evaluation and peer ratings. Cashin (1995) reports that classes in which students give the instructor higher ratings tend to be the classes in which students learn more; and that moderate to high correlations were found between student ratings and administrator, colleague, and alumni ratings. In the literature, student ratings are viewed as both reliable and valid.

The major criticism of using student evaluation forms to collect feedback about teaching effectiveness is the potential for the results to be affected by biases related to characteristics of the instructor, the student, the course, and/or the data collection method. Much quantitative research has been conducted in this area. Variables that are significantly related to student ratings include: faculty rank, instructor expressiveness; student motivation, expected grades, and gender; level of course, academic field, workload difficulty, and elective versus required courses. The presence of the instructor while the survey is administered and the requirement that the student sign the form also significantly affect the ratings. Variables with little or no relationship to student ratings include: instructor age, teaching experience, gender, race, personality and research productivity; student age, level, GPA, and personality; class size and time of day; and the time data are collected during the term.

Student ratings are well researched and the results support their reliability and validity; but they remain controversial and widely criticized by faculty members. Faculty question the appropriateness of being judged by students, doubt the validity of the evaluations, and express other reasons for discrediting the results.

Faculty Evaluation Programs

As one reads the writings of different authors on the subject of faculty evaluation, "one is struck by the high degree of agreement among them. ... there is 80 to 90 percent agreement about the general principles that should guide effective faculty evaluation" (Cashin, 1996, p.1) Writers on faculty evaluation almost unanimously recommend the use of multiple sources of data. No single source of data – including student ratings – provides sufficient information to make a valid judgement about overall teaching effectiveness (Braskamp & Ory, 1994; Centra, 1993; McKeachie, 1999; Theall & Franklin, 2000).

Accountability demands and faculty development needs call for different evaluation methods and procedures (Johnson & Ryan, 2000, p.117). Unfortunately, in most institutions, more emphasis is placed on students' evaluation of teaching and not enough on other methods. At many educational institutions, the sole method for the evaluation of teaching is student ratings. The majority of authors reviewed recommend the use of multiple sources of data to evaluate teaching effectiveness; and the avoidance of making any decisions about an instructor's teaching

effectiveness based on data from a single source. Johnson and Ryan (2000) present six teaching evaluation methods which include: student ratings; self-evaluations; teaching portfolios; instructor/student interviews; alumni reports and peer reviews. Seldin (1999) describes 15 different sources for information (see Table 1) that can be used to evaluate an instructor's effectiveness. Research is rigorously evaluated by peers yet academics resist having teaching evaluated in the same rigorous manner.

Effective Teaching in Higher Education

One assumption about teaching in higher education is that students are adult learners or are in the process of becoming adult learners. First year students need teaching methods that support this development. Such methods differ from those that are appropriate for fourth year students. Principles of adult learning are useful in assisting in the design and conduct of quality courses in higher education.

While no single list of teaching qualities has been developed that satisfies all interested stakeholder groups, the numerous lists that can be found in the literature imply seven common characteristics:

1. enthusiasm for the subject;
2. passion for teaching;
3. strong communication skills;
4. subject area expertise;
5. empathy for the learner;
6. understanding of the teaching and learning process; and
7. understanding of alternative learning styles and teaching methods.

According to Cranton (1998): "We need to know about teaching and learning, to come to understand ourselves as educators, then relate these understandings to our discipline and context, and thereby develop our own practice" (p.48). A recent development in the assessment of teaching is a focus on the concept of the scholarship of teaching. Boyer (1990) defines faculty work as calling for four types of scholarship: the scholarship of discovery, integration, application, and teaching. This conceptual model has generated much interest and considerable recent research. As a result, institutional and faculty perceptions of teaching are being challenged and are changing.

One of these changes is a shift from an emphasis on teaching to an emphasis on learning as the major focus on any classroom (Barr & Tagg, 1995). To be rigorous in the evaluation of teaching therefore requires a fundamental change in approach – one that shifts the focus of evaluation from surface features to deeper structures (Biggs, 1999; Pratt, 1997). A common thread running through the literature on the scholarship of teaching is the significance of reflection as an essential feature in improving the quality of teaching in higher education. This study contributes to the emerging body of literature on the scholarship of teaching and learning, and illustrates the need for further study on the role of evaluative feedback in both teaching and the scholarship of teaching.

Methodology

A qualitative phenomenological case study approach was chosen to explore a phenomenon — the evaluation of teaching. The collective case study is ‘phenomenological’ because it examines the lived experiences of a small number of participants (26) through extensive interviewing to develop patterns and relationships of meaning (Creswell, 1994, p.12). This study presents highlights from the data gathered from five research sites, illustrating the research participants’ perceptions of the formal student evaluation processes, other formal and informal teaching evaluation practices, and their reflections on teaching.

In-depth interviews were conducted to gather data from the unit head and up to five faculty members at each of five schools of business in Atlantic Canada. Six research questions guided this study:

1. How do professors perceive formal university procedures used to gather information that evaluates teaching?
2. What are the procedures at their university?
3. How do professors use information gathered through formal evaluation procedures?
4. How do professors informally gather information pertaining to their teaching performance? How do they use such information?
5. How do professors perceive the university to use formal and informal assessment data to make decisions about tenure and promotion?
6. How do professors perceive the university to use formally and informally gathered information on teaching evaluation to encourage better teaching?

Discussion of Findings and Themes

Two fundamental needs were present throughout the data: the individual faculty member’s need to improve the quality of her or his teaching; and the institution’s need to assess the quality of individual teaching. Five major themes emerged:

1. the use of student ratings as the primary (and sometimes only) source of data used to assess teaching;
2. the need to use multiple sources and types of information to assess the quality of teaching;
3. the role of reflection, self-reflection and self-assessment in the improvement of teaching;
4. the need for continuous improvement in the quality of post-tenure teaching; and
5. the equitable assessment of teaching and research.

These five themes illustrated components of both of the overarching themes, although often in conflicting ways. Following is a discussion of the findings and themes.

Theme One:

The use of student ratings as the major, and often only, source of information in faculty assessment procedures is problematic.

The participants in this study make little use of feedback from the statistical summaries of student ratings for three basic reasons:

- Timing, the feedback -- whether used for formative or summative purposes -- is not received in a timely fashion or often enough. The results usually are not available to make changes in subsequent courses.
- Credibility -- the student ratings are perceived as being easily manipulated through the administration procedures. The results are perceived as being biased.
- Unhelpful -- the quantitative questions asked on the student ratings forms are perceived as the wrong questions; and the answers, therefore, as having little value to the instructor. The only questions that are commonly used by the participants to gauge the quality of their teaching are the global questions that ask: *In your opinion was this a good course?* and *In your opinion, is this instructor an effective teacher?*

A further concern for the participants was that the statistical summaries can be interpreted in many different ways making the use of the data for important career decisions questionable and often unfair. The only aspect of the student ratings procedure that was viewed positively was the students' responses to open-ended questions. Instructors rely on handwritten student responses to open-ended questions to make improvements in their teaching and to improve students' learning.

In quality of teaching terms, institutions use student ratings to recognize and reward faculty members with the highest ratings and to encourage improvement for those with the lowest ratings. From the faculty member's perspective, quantitative ratings are problematic because, whether recognized or encouraged to improve, they do not receive any accompanying information about how they could improve their teaching. From the institution's perspective, the quantitative results can be used to justify personnel decisions and to report, in simplified terms, on the quality of teaching achieved in classrooms.

These differences help explain the conflict that appears to exist between instructor and institutional perceptions of student ratings. The institutional perception is based on the quantitative or positivistic approach, using summative data to draw conclusions about teaching quality. Student surveys are easy to administer and cost effective to analyse; the resulting data are useful in making promotion and tenure decisions and provide a relatively simple way to support claims about quality of teaching within the institution. Faculty perceptions are based on perceived biases related to administering the surveys in class, on the lack of meaning in the statistical analyses, and on the lack of information that might lead to improvements in teaching. Further, most of them perceive that the quality of their teaching is not as important to further their careers as the quality and quantity of their research.

Theme Two: The assessment of teaching must be based on multiple sources and multiple types of information to ensure equity for the faculty member and quality for the institution.

The reliance on a single source of data to make important decisions that affect a faculty member's career is perceived by study participants as both inadequate and unfair. Each participant used additional sources and means to obtain feedback about their own teaching.

These means included asking students directly for feedback, either in writing or orally; reviewing how well students perform on tests and examinations; and discussing concerns with peers. The participants believed that any assessment of the quality of teaching should be based on more than one source and one type of data; that multiple sources and multiple types of data would ensure a more equitable evaluation process. While the question was not asked of the institutions that housed the five research sites, the literature indicates that the use of multiple sources and multiple types of data also ensures a better indication of the quality of teaching that can be found within its classrooms (Cashin, 1996; Ory, 2000; Pratt; 1997; Seldin, 1999).

Institutions could require candidates being considered for promotion and tenure to submit multiple sources of information about the quality of their teaching. An increased emphasis on the evaluation of teaching performance would be a good thing for everyone involved — the instructor, the students, the faculty and the institution as a whole. By raising the standards and expectations for university teaching, institutions could foster a corporate culture based on excellence in teaching. Lunde and Barrett (1996) suggest that “building evaluation of teaching into the fabric of the faculty’s personnel processes guarantees that teaching will have importance comparable to that of research and other scholarly activities” (p.96). Two additional sources of assessment data most frequently discussed by the participants were peer observations and alumni ratings of teaching.

Theme Three: Reflection, self-reflection and self-assessment — for example, through classroom-based research — are key components in the process of improving teaching.

Instructors need to rely on self-reflection and on valid feedback from others when assessing their own teaching performance. Reflection in academia involves thinking about teaching through such activities as reading and engaging in discussions with others. Self-reflection involves instructors in a regular review of the processes they use in their classes stimulated by feedback from students, peers and administrators; and of student performance on tests and assignments. Self-assessment combines reflection and self-reflection to assess teaching and learning effectiveness by comparing one’s own performance with standards espoused by others and with criteria established by self. Institutions need to provide opportunities for instructors to engage in these reflective activities. Such opportunities require a commitment in terms of time as well as supportive resources.

Most participants in this study were being encouraged to reflect on their teaching once a year through reviewing student ratings or writing a self-report. Most self-reports cover research, publications, and service as well as teaching activities. In only a few cases does this self-report also include evaluative statements about quality of teaching. Pre-tenured faculty members at all sites were required to write an annual self-report and to meet with the unit head to discuss their performance and career development plans.

The literature on the quality of teaching stresses the importance of reflection, self-reflection and self-assessment to inform instructional behaviours and facilitate change (Kreber & Cranton, 2000; Schön, 1995; Smith, 2001a; Trigwell et al., 2000). As a source of information used to evaluate teaching in universities, self-evaluation has gained popularity in recent years (Seldin, 1999). Reflection and self-evaluation play a significant part in a multifaceted evaluation process.

Sorcinelli (1999) describes a standard teaching self-evaluation form used by some institutions for annual reviews. Unlike a faculty activity sheet that calls for listing all activities (teaching, research, publications, and service), the teaching self-evaluation form focuses only on teaching activities. Such forms not only ask what individuals did but also how well they think they did. Faculty are asked to rate the frequency of their specific teaching behaviours on a scale of one to five. The benefit of such forms is that a similar form could be used to ask students to rate the same behaviours. The faculty member would then be able to compare self-ratings to student ratings.

Theme Four: The granting of tenure should not end the requirement for improvement in quality of teaching

Maintaining excellence in an academic institution requires continuous development of its most critical resource: its tenured faculty. This necessitates a constructive developmental program of post-tenure review. Once tenured, senior faculty are usually left to set their own professional quality and performance levels. Also, there is little incentive for senior faculty to continue to improve the quality of their teaching as they might have been encouraged before receiving tenure.

Much discussion about post-tenure review is found in the literature. Some North American institutions are now requiring that post-tenured faculty go through a performance review similar to that experienced by pre-tenured faculty (Alstete, 2000; Licata & Morreale, 2001). Some institutions choose to review annually; others conduct performance reviews every five years. Such reviews are normally conducted by a committee of peers and administrators. These reviews are designed to encourage faculty to continuously improve or at least to sustain their teaching productivity and could assess both course design and teaching performance.

The participants in this study perceive that poor teaching by post-tenured faculty members is either ignored or little action is taken to force an instructor who consistently receives poor student ratings to improve the quality of their teaching. Such faculty members are unlikely to request feedback about their teaching by any other method or from any other source. Until collective agreements with faculty are changed, this situation will continue.

Theme Five: Teaching and research are not treated equally in faculty promotion and tenure procedures and decisions.

The rhetoric in academe is that professors should devote 40% of their efforts to each of teaching and research, with a service component accounting for the remaining 20%. Most participants in this study thought that teaching should be considered as equal to research in importance. However, the majority also thought that their institutions, including those that pride themselves on being teaching institutions, were moving in the direction of placing greater emphasis on research. The literature suggests faculty may choose to spend more time thinking and reflecting on their research because reward structures in academia are set up to reward good research more easily than good teaching (Higgerson, 1999; Svinicki & Menges, 1996).

Grants for research are more abundant than grants for teaching. The only teaching grants available at the participating sites were for planning and implementing new courses or for innovative teaching initiatives usually involving technology. While good teaching was viewed as necessary by all participants, they all believed that tenure and promotion decisions placed heavy emphasis on research and publication records.

Teaching and research are touted as being equal in their significance to a university; however, the two activities are not perceived as being treated equally during tenure and promotion decisions. Efforts to improve teaching must be rewarded in ways that are equitable with the ways in which research is rewarded. Research is rewarded through research awards, reduced teaching load, recognition for related publications, and higher recognition in promotion and tenure decisions. Teaching could be rewarded through teaching awards, reduced teaching load, recognition for scholarly teaching publications, and equal recognition in promotion and tenure decisions. Teaching awards, however, are perceived by the participants as merely popularity contests with nomination and selection procedures heavily criticized by the participants. Institutions have many effective teachers but only recognize a few. The actual awards are not very substantial with the public recognition of teaching excellence meaning more to the recipient than the monetary value of the award.

Conclusion

The findings of the collective case study have been categorized into five major themes and the relationship between these themes and the two fundamental needs: the faculty member's need to improve the quality of his or her teaching and the institution's need to assess the quality of individual teaching, is illustrated in Table 2.

“Many student-based evaluation procedures aimed at evaluating the teacher and teaching methods seem out of date and out of touch” (Braskamp, 2000, p.27). The data from this study support opinions expressed in the literature that teaching evaluation processes should rely on multiple types of information from multiple sources, particularly when others are making judgments about an instructor's teaching competence (Cashin, 1995; Seldin, 1999). Multiple sources of information are needed for a full and equitable view of an instructor's quality of teaching. When assessing the quality of an instructor's teaching, however, most institutions rely heavily on the ratings obtained from a formal student evaluation process. Additional sources of information could include: written comments from students; peer observation data; self-evaluation assessments and reports; alumni data; and information from department heads and deans.

References

- Alstete, J.W. (2000) Post tenure faculty development: Building a system for faculty improvement and appreciation. *ASHE-ERIC Higher Education Report*, 27 (4), 8-11.
- Angelo, T.A. & Cross, K.P. (1993) *Classroom assessment techniques: A handbook for college teachers* (2nd edition). San Francisco: Jossey-Bass.
- Barr, R.B. & Tagg, J. (1995) From teaching to learning – A new paradigm for undergraduate education. *Change*, 27(6), 12-25.
- Biggs, J. (1999) *Teaching for quality learning at university*. Buckingham, UK: Open University Press.
- Boyer, E.L. (1990) *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: Princeton University Press.
- Braskamp, L.A. (2000) Towards a more holistic approach to assessing faculty as teachers. In K. Ryan (Ed.), *Evaluating teaching in higher education: A vision for the future*. New Directions for Teaching and Learning (no. 83) (pp.19-33). San Francisco: Jossey Bass.
- Braskamp, L. & Ory, J.C. (1994) *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco: Jossey-Bass.
- Cashin, W.E. (1996) *Developing an effective faculty evaluation system*. IDEA Paper No. 33. Manhattan, KS: Kansas State University, Centre for Faculty Evaluation and Development. ERIC Reproduction Document ED 395 536.
- Cashin, W.E. (1995) *Student ratings of teaching: The research revisited*. IDEA Paper No. 32, Manhattan, KS: Kansas State University, Centre for Faculty Evaluation and Development. ERIC Reproduction Document ED 402 338.
- Centra, J.A. (1993) *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass.
- Council of Ministers of Education, Canada (CMEC). (1999) *A report on public expectations of postsecondary education in Canada*. Toronto: CMEC. Retrieved 03/28/2000, from www.cmec.ca/postsec/expectations.en.pdf
- Cranton, P. (1998) *No one way: Teaching and learning in higher education*. Toronto: Wall & Emerson.
- Creswell, J. W. (1994) *Research design: Qualitative & quantitative approaches*. Thousand Oaks, CA: Sage.
- Feldman, K. (1997) Identifying exemplary teachers and teaching: Evidence from student ratings. In R.P. Perry & J.C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 368-395). Edison, NJ: Agathon.
- Glassick, C.E., Huber, M.T. & Maeroff, G.I. (1997) *Scholarship assessed: Evaluation of the professoriate*. San Francisco: Jossey-Bass.
- Higgerson, M.L. (1999) Building a climate conducive to effective teaching evaluation. In P. Seldin & Associates, *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions* (pp. 194-212). Bolton, MA: Anker Publishing.
- Isaacson, R. (2000) Why JoSoTL and why now? *The Journal of Scholarship of Teaching and Learning*, 1 (1), 2-7. An electronic journal accessed through <http://www.iusb.edu/~jostl/>

- Johnston, T.D. & Ryan, K.E. (2000) A comprehensive approach to the evaluation of college teaching. In K. Ryan (Ed.), *Evaluating teaching in higher education: A vision for the future*. New Directions for Teaching and Learning (no. 83) (pp. 109-123). San Francisco: Jossey-Bass.
- Knapper, C. & Cranton, P. (Eds.) (2001) *Fresh approaches to the evaluation of teaching*. New Directions for Teaching and Learning (no. 88). San Francisco: Jossey-Bass.
- Kreber, C. & Cranton, P. (2000) Exploring the scholarship of teaching. *Journal of Higher Education*, 71(4), 476-495.
- Licata, C.M. & Morreale, J.C. (2001) Implementing post-tenure review. *American Association for Higher Education Bulletin*, December. Accessed on 10/12/2002, from <http://aahebulletin.com/public/archive/implementingptr.asp>
- Lunde, J.P. & Barrett, L.A. (1996) Decentralized/departmental reward system. In M.D. Svinicki & R.J. Menges (Eds.), *Honoring exemplary teaching*. New Directions for Teaching and Learning (no. 65) (pp. 93-98). San Francisco: Jossey-Bass.
- Marsh, H.W. & Dunkin, M.J. (1997) Students' evaluations of university teaching: A multidimensional perspective. In R.P. Perry & J.C. Smart (Eds.), *Effective teaching in higher education: Research and practice* (pp. 241-319). Edison, NJ: Agathon.
- McKeachie, W. J. (1999) *Teaching tips: Strategies, research, and theory for college and university teachers* (10th edition). Boston: D.C. Houghton Mifflin.
- Murray, H. G. (1997a) Does evaluation of teaching lead to improvement of teaching? *International Journal of Development*, 2 (1), 8-23.
- Ory, J.C. (2000) Teaching evaluation: Past, present, and future. In K. Ryan (Ed.), *Evaluating teaching in higher education: A vision for the future*. New Directions for Teaching and Learning (no. 83) (pp. 13-18). San Francisco: Jossey-Bass.
- Pratt, D.D. (1997) Reconceptualizing the evaluation of teaching in higher education. *Higher Education*, 34, 23-44.
- Schön, D. (1995) The new scholarship requires a new epistemology. *Change*, 27 (6), 26-34.
- Seldin, P. (1999) Current practices: Good and bad – nationally. In P. Seldin & Associates, *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions* (pp. 1-24). Bolton, MA: Anker Publishing.
- Smith, R. (2001a) Formative evaluation and the scholarship of teaching and learning. In C. Knapper & P. Cranton, (Eds.), *Fresh approaches to the evaluation of teaching*. New Directions for Teaching and Learning (no. 88) (pp. 51-62). San Francisco: Jossey-Bass.
- Sorcinelli, M.D. (1999) The evaluation of teaching. In B. Pescosolido & R. Aminzade (Eds.), *The social worlds of higher education: Handbook for teaching in a new century*. Thousand Oaks: Pine Forge Press.
- Svinicki, M.D. & Menges, R.J. (1996) Consistency withing diversity: Guidelines for programs to honor exemplary teaching. In M.D. Svinicki and R.J. Menges (Eds.), *Honoring exemplary teaching*. New Directions for Teaching and Learning (no. 65) (pp. 109-113). San Francisco: Jossey Bass.
- Theall, M. & Centra, J. (2001) Assessing the scholarship of teaching: Valid decisions from valid evidence. In C. Kreber (Ed.), *Scholarship revisited: Perspectives on the scholarship of teaching*. New Directions for Teaching and Learning (no. 86) (pp. 31-44). San Francisco: Jossey-Bass.
- Theall, M. & Franklin, J. (2000) Creating responsive student ratings systems to improve evaluation practice. In K. Ryan (Ed.), *Evaluating teaching in higher education: A vision*

for the future. New Directions for Teaching and Learning (no. 83) (pp. 95-107). San Francisco: Jossey Bass.

Trigwell, K., Martin, E., Benjamin, J. & Prosser, M. (2000) Scholarship of teaching: A model. *Higher Education Research & Development*, 19 (2), 155-168.

Table 1
Information Sources ‘Always Used’ in Evaluating Teaching Performance
10 Year Comparison (in %)

<i>Information source</i>	<i>1988</i> (n = 604)	<i>1998</i> (n = 598)	
Systematic student ratings	80.3	88.1	+
Evaluation by department chair	80.9	70.4	-
Evaluation by dean	72.6	64.9	-
Self-evaluation / report	49.3	58.7	+
Committee evaluation	49.3	46.0	-
Colleagues’ opinions	44.3	44.0	-
Classroom visits	27.4	40.3	+
Course syllabi and exam review	29.0	38.6	+
Scholar research / publication	29.0	26.9	-
Informal student opinions	11.3	15.9	+
Alumni opinions	3.0	9.0	+
Grade distribution	4.2	6.7	+
Long-term follow-up of students	3.2	6.0	+
Student examination performance	3.6	5.0	+
Enrollment in elective courses	1.2	1.5	+

Adopted from: Seldin, P. (1999) Current Practices: Good and Bad - Nationally. In P. Seldin and Associates, *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions* (Table 1.5, p.14). Bolton, MA: Anker Publishing.

Table 2
Relationship of Fundamental Needs and Major Themes

Major Themes	Individual Instructor Needs to Improve Quality of Teaching	Institutional Need to Assess Quality of Teaching
<p>Theme One</p> <p><i>Use of student ratings as primary source of data to assess quality of teaching</i></p>	<p>Student ratings perceived by faculty member as:</p> <ul style="list-style-type: none"> • unreliable and biased • provide information that does not help improve teaching 	<p>Student ratings perceived by administrators as:</p> <ul style="list-style-type: none"> • easy to administer and cost effective • provide simplified form of evidence for quality of teaching in personnel decisions and for publicity information
<p>Theme Two</p> <p><i>Use of multiple sources of information to assess quality of teaching</i></p>	<p>To improve quality of teaching, faculty members perceive:</p> <ul style="list-style-type: none"> • quantitative data as not providing useful feedback • qualitative data, informal assessments, and peer observations as providing useful feedback 	<p>To assess quality of teaching administrators perceive:</p> <ul style="list-style-type: none"> • qualitative data as unreliable and too subjective to be used in personnel decisions • qualitative data can be used to make awards for teaching excellence • no consistent criteria or means for using qualitative data to assess teaching • no well-trained peer observers who could do peer observations
<p>Theme Three</p> <p><i>Reflection, Self-reflection and self-assessment</i></p>	<ul style="list-style-type: none"> • done formally and systematically with pre-tenured faculty • no funds available to support classroom-based research • few opportunities provided to support systematic reflection and self-assessment for post-tenured faculty 	<ul style="list-style-type: none"> • self-assessments not seen as reliable or valid • no consistent criteria or means for using self-assessments in quality of teaching decisions • do not support or encourage classroom-based research • funds provided to develop new courses or to revise existing course for use with new technologies
<p>Theme Four</p> <p><i>Ongoing improvement in quality of teaching following tenure</i></p>	<ul style="list-style-type: none"> • strong concern about teaching in early years of career, diminishes with experience and tenure • concerns about research and concerns about teaching tend to distract each other 	<ul style="list-style-type: none"> • concern about research high throughout career • concern about teaching only when consistently poor student ratings are received • faculty receiving poor student ratings cannot be forced to improve their teaching
<p>Theme Five</p> <p><i>Assessment of teaching and research inequitable</i></p>	<ul style="list-style-type: none"> • need to learn how to use institution-established criteria for assessing teaching, research and service 	<ul style="list-style-type: none"> • need to establish clear criteria for assessing teaching, research and service

Summary

This paper reports the findings of a qualitative study of faculty members' perceptions of teaching evaluation processes to understand how information from student ratings and other instruments were used to assess teaching behaviours in higher education institutions.

In-depth interviews were conducted to gather data from the unit head and up to five faculty members at each of five schools of business in Atlantic Canada. All sites used student ratings as the primary source for information about teaching, required a self-assessment of pre-tenured faculty, and annually recognized teaching excellence. No site had a functioning peer evaluation program for the improvement of teaching. Information gathered through the formal student survey process was used by the institutions as an indicator of teaching effectiveness in relation to applications for tenure, promotions, or contract renewal. The student survey results were included in a faculty member's personnel or official file, and could be used in faculty-developed teaching portfolios.

All participants viewed the student survey process as unhelpful to bring about change in their teaching. Feedback from student ratings was often too late to be useful, given a cursory review, and quickly filed away. They viewed student ratings as easy to manipulate and easy to misinterpret. Nearly all participants sought informal feedback on their teaching from students with the majority directly asking students. Student outcomes, if assessed carefully and used cautiously, were perceived helpful in evaluating teaching, however, student outcomes were also perceived as extraordinarily difficult to correlate with specific interventions. Qualitative data, obtained from open-ended questions on the student survey and through informal means, were seen as helpful to improve teaching.

The heavy emphasis on the feedback from the student opinion surveys to make inferences about an individual's teaching ability was perceived by the participants as inappropriate. Participants felt little pressure to improve their teaching because poor teaching, as measured by the formal student evaluation process, was often simply ignored. Teaching excellence was recognized through annual awards to motivate quality in teaching, however, these awards were viewed with some skepticism because such details as the nomination and selection processes were unclear.

Five major themes emerged from the data: 1. the use of student ratings as the primary (and sometimes only) source of data used to assess teaching; 2. the need to use multiple sources and types of information to assess the quality of teaching; 3. the role of reflection, self-reflection and self-assessment in the improvement of teaching; 4. the need for continuous improvement in the quality of post-tenure teaching; and 5. the need for equitable assessment of teaching and research. Two fundamental needs were illustrated in these themes, although often in conflicting ways: the individual faculty member's need to improve the quality of her or his teaching; and the institution's need to assess the quality of individual teaching.