

## Introduction

To an outsider, the decisions at a university about who receives promotions, merit increases, and tenured positions are shrouded in mystery. The process, even to those of us on the inside, is complicated and needs the careful guidance of the Faculty Association and the University's administration. They, as well as the Canadian Association of University Teachers, caution professors involved in peer review to exercise "honesty and fairness and integrity", reminding us that the deliberations of a peer review committee are part of a collegial process. What this means is that one must gather and weigh evidence and make recommendations based on that evidence.

What kind of evidence can be gathered by observing a colleague teaching a class?

## Some Issues

While its relative importance varies from Faculty to Faculty at this and most universities, teaching is a major factor in assessing the overall performance of a faculty member. The techniques and criteria for such assessment of teacher effectiveness will remain controversial, in part because determining the relationship between teaching and learning remains problematic. If no learning occurs how can we say that there was any teaching done? Of course much learning can happen even when there is no effective teaching. It is also fair to say that those faculty who receive positive evaluations from students and colleagues believe those evaluations to be fair and valid while those who receive negative evaluations from the same sources dismiss this kind of information as less than significant.

## Types of Evidence

John Centra, one of the American leaders in the field of the evaluation of faculty, has identified fifteen different methods of evaluating teaching in Universities in North America:

1. systematic student ratings
2. informal student opinions
3. colleague ratings based on classroom visits,
4. colleague opinions based on "other" evidence,
5. student performance on exams,
6. peer evaluation of course outlines and exams,
7. department chairperson's evaluation,
8. dean's evaluation
9. committee evaluation,
10. self-evaluation,
11. long-term follow-up of students' performance
12. alumni opinions or ratings
13. level of student enrolment in instructor's elective courses,
14. analysis of videotapes of classroom instruction, and
15. participation in teaching improvement activities such as work-shops and courses.

## **Determining Quality**

Centra (1993) found that chairpersons' evaluations, systematic student ratings, and colleagues' opinions were the most frequently used and influential methods of evaluating teaching for promotion, tenure, and salary purposes. Often these methods are simply the most convenient and cost effective in terms of people's time and the limited resources of the department. Faculty and administrators, often those who sit on peer review committees, are often implicitly presumed to be able to judge the quality of a university teacher's effectiveness. Problems tend to arise when chair's evaluations and colleagues' opinions are formulated without observing for themselves what goes on in the classroom of the professor under review.

Assessing what constitutes effective teaching remains problematic when there is only one form evidence, which in most cases is likely to be student evaluations of performance. That is why classroom observations by peers is seen as a broadening complement to student evaluations. We also know that students are very good at providing opinions on some teacher behaviours, like clarity and rapport for example, but less adept at assessing such things as instructor knowledge and course content.

## **What Colleagues Can Do**

Even without entering a class, colleagues in the same discipline or department should be able to assess:

1. the instructor's knowledge of the subject,
2. selection of course objectives and pertinence of the reading list
3. scope and sequence of the syllabus,
4. appropriateness of instructional materials, projects, assignments and exams.
5. supervision of student research and theses,
6. involvement in teaching development activities,
7. curriculum development and preparation of instructional materials, and
8. research and scholarship related to teaching.

Our colleagues are much more comfortable reviewing their peers' teaching documents than observing them in the classroom (Braskamp and Ory, p. 97). This says as much about the print dependent culture of academics as it does about a lack of training in what and how to observe in a classroom or lecture theatre.

## **Classroom Observations**

Many in the field of faculty evaluation have noted, however, that classroom visits and observations remain one of the most controversial of the methods of evaluating teaching in higher education. There are many reasons for this, including an academic culture that has traditionally not regarded teaching as a shared or observable activity. In part because professors have had little or no preparation for teaching, professors may display the defensiveness of those who have been self-taught largely through trial and error.

Sometimes this translates into cries that classroom visits by peers are an infringement of

academic freedom or that they are demeaning to a person's standing in the department. Unfortunately classroom observations often only occur when a "problem" has been identified through student ratings or complaints or exam results. This can lead to classroom visits being seen as punitive rather than formative.

## **A Critical Perspective**

Robert Menges argues (in Seldin, 1984, p. 112) that reports by colleagues of classroom visits will remain poor sources of data unless there is a "coherent and sound program of many visits over time by several colleagues who have been trained to produce reliable reports". This is simply not practical for most departments.

It will take a long time to change the culture surrounding university teaching so it becomes a topic of collaboration as well as of scholarly action research. However it is noteworthy that newly-appointed professors evince a more open attitude than tenured faculty to sharing their teaching and ideas about teaching and are more willing to engage in an on-going process of improving their teaching.

Centra, and Braskamp and Ory (1994), among others, are sceptical of the efficacy of collegial classroom visits for three reasons. The reliability of the data from such visits remains in question because colleagues tend to bring biases to their observations, have not been trained to observe adequately, and are often inconsistent in what they do or do not observe. Seldin (1984) too provides the codicil that collegial visits should be encouraged in order to provide formative feed-back for improvement rather than summative assessment for personnel decisions.

The literature does suggest that most colleagues, for whatever reasons, rate their peers as good or even excellent instructors (Dilts et al, p. 58). While colleagues may readily observe a peer's "performance" (or surface characteristics) to which they can relate because they've "been there themselves", they are less able to analyse the deeper structure of a lesson or lecture, much less determine whether or not there is a connection between the observed teaching behaviours and student learning. This is especially so if the lecture has not been structured to allow for feed-back or interaction that demonstrates student engagement with the material.

## **The Emphasis on Process**

Seldin (1980) makes the point that both those being observed and those doing the observing must be in agreement about what should be looked for. This implies a meeting between the colleagues before the class is observed to reach consensus on what is to occur. This process is as important as the particular data gathering instrument used. Teacher suspicion and resistance and observer's mishandling of the instrument will nullify the process. The success or failure of the peer observation process really depends of two things: the spirit of collegiality, and the shared perception that the rating instrument used is fair and appropriate. Ensuring that these two elements are in place is a primary and sometimes onerous responsibility of Heads of Departments and Deans of Faculty.

## **How are Things Changing?**

Over the past thirty years, Kenneth Eble has seen a steady move in peer evaluation to “more care in gathering data, more attention to ruling out prejudice and subjectivity, and more involvement of those actually affected by the process” (in Seldin 1984, p. 96). The “clubby” atmosphere of the fifties is long gone. Eble sees many reasons for the changes in the peer evaluation process including increased accountability to students and the public, the loss of collegiality and trust within departments, and “the fear of grievances and litigation”(p. 97).

“From monastery to industry” is the phrase George Geis uses to describe changes in university life (Seldin, 1984, p. 101). The image of an isolated scholar committed primarily to his own learning and through it to his discipline is an archaic one. The industrial model takes students, money and research as raw materials and then processes them into clients, products and out-puts.) As the life of a professor has become a unionised job and no longer a calling to the ivory tower, evaluation has become more explicit, more accountable, and more dependent on data. Geis anticipates greater effort being made to evaluate entire programs and departments in light of the goals of the institutions because this is a legitimate way to gauge the validity of the data gained through peer evaluation. In other words, peer evaluation is less problematic if the entire university community agrees that professors are primarily teaching scholars.

## **Conclusion**

“All the litigation, and even its threat, by professors denied promotion or tenure is forcing institutions to take special pains to find hard data on which to base their decisions”. ( Seldin 1984, p. 116) Yes, we need “honesty, fairness and integrity” in the process but it would be naive to believe that pettiness, internal political, ideological, and personality conflicts will be banished from promotion/tenure decisions. But if colleagues, in assessing the teaching of their peers, can agree on the appropriateness of the data gathering observation process and instruments, they will have taken an important first step in assessing and improving teaching in higher education.