

## CHAPTER 6

### THE CONTEXT FOR EVALUATION AND PROFESSIONAL DEVELOPMENT AT THE COLLEGE AND DEPARTMENTAL LEVELS

Christie Evans worked for the College of Arts and Sciences at Midwestern University. This was the largest college on campus and offered majors and specialties in all the disciplines from the natural and physical sciences, the humanities to the social and behavioral sciences.

The college represented in small scale the characteristics of the university. It comprised 25 % of the faculty and approximately 38% of the student population at Midwestern. Approximately, half of the students were males and half were females. Most of the students were White students (70%). Minority students comprised 26% of the student population, and 4 % of the students were international students. Of the 609 faculty members, eighty six percent were White. Sixty-nine percent of the faculty were men and thirty one percent were women.

The college nourished a reputation for "good teaching," noting the number of teaching awards obtained by its faculty. The College also created several awards to reward good teaching in addition to those provided at the college and departmental levels. As a professor mentioned:

All of the campus and college awards for teaching show that the university cares about teaching. Before I came to this university I taught at Yale, but there what mattered the most was scholarship. Here, the university has a system to monetarily reward good teaching. There is an award for being a good instructor that is about \$5,000. It can be for undergraduate, graduate teaching, the introduction of innovations, or the use of technology in teaching.

According to one of the College's Associate Deans, college support for teaching was influenced by leadership at the college and, especially, at the departmental level:

In order for the departments to value teaching, leadership at the college and unit levels is important. The Dean of the College of Arts and Sciences is very interested in the improvement of teaching. So, there is support at the college level but leadership at the departmental level is vital because the improvement of teaching has always been a responsibility of the departments.

The college was said to be open to different views of teaching and different ways for evaluating teaching. According to the Associate Dean, the only aspect that was stressed by the administration of the college to its departments was the importance of linking teaching to student learning:

There is reciprocity between teaching and learning. Good teaching is linked to learning; it is a reciprocal process. As a college, we encourage instructors to see and foster this relationship. Nowadays, we know that the person who is learning is also important. So, you have to focus on learning but this doesn't imply that you have to teach in a certain way. Good teaching happens, no matter what it looks like. Good teaching is situational. If a student comes out of a class with a new conception, a new understanding of the content, good teaching took place.

The Associate Dean added that good teaching was reflected in the global influence of the instructor on students' learning:

Good teaching is global more than multidimensional. It involves the overall influence of the teaching on students' learning. Good teaching involves those intangibles that engage students with the subject. For example, you can have someone who is not very good at formal lecturing, but who because of other elements such as the delivery or the activities in the classroom, conveys an attachment of passion for the subject that help students learn.

The evaluation of teaching was to be conducted according to campus guidelines for tenure and promotion as well as for the college policies on teaching. Departments and college units had the main responsibility for the evaluation of teaching.

#### College Policies and the Evaluation and Improvement of Teaching

Campus policies for evaluating teaching were followed in the several units of the College of Arts and Sciences. There was no conflict between campus and college legislation regarding the evaluation of teaching.

The College's Policy Manual supported in-depth evaluation of teaching and provided a list of guidelines for evaluating teaching that were consistent with campus legislation. The manual also was consistent with the Provost's communications on tenure and promotion regarding not including student comments from Teaching Evaluation System (TES) forms for making administrative decisions. Both the manual and the Provost's communications for tenure and promotion stressed the need for the evaluation to be based on evidence, not on hearsay. Both documents stressed the importance of student ratings in the evaluation of teaching. Even though no research had been conducted on campus about the real meaning of the scores and their relevance and appropriateness for making administrative decisions about teaching, campus and college legislation considered the ratings as a reliable evidence for evaluating instruction.

Campus legislation gave higher importance to the use of TES ratings<sup>1</sup> over peer evaluations. On the other hand, the College of Arts and Science's Policy Manual put more

---

<sup>1</sup> According to the Provost communication for promotion and tenure: "Teaching evaluation **must** include a summary of TES data (or a summary developed through use of a departmental instrument, in the

emphasis on the use of peer evaluation but stressed that classroom observations should not include evaluative statements:

Experience to date suggests that the most useful evaluative statement is one prepared by a team of two or three teachers who are able to examine the reviewee's teaching materials and have attended some classes. This mode of evaluation, already used in some departments, is encouraged further. Peer observations are meant to be simple descriptions (without evaluation) of the candidate activities." (College Policy Manual, 1999, p. 7)

According to one of the college's the Associate Deans, the emphasis that campus legislation put on the use of TES ratings and the way in which this was applied, it was said to negatively influence the evaluation of teaching:

Presently, teaching has increased its role in promotion and tenure. Campus guidelines for assessing teaching during tenure and promotion are very strict. This legislation requires professors to use the TES scores for at least two semesters. Most committee members also emphasize the use of TES scores over other sources of information. As a result, important contextual information may not be included. They like the scores because they are numbers that they can graph.

#### Teaching Evaluation and Professional Development at the College Level

Decisions about teaching and mentoring were the responsibility of the departments. Departments had the responsibility for implementing mechanisms for evaluating and making decisions about teaching.

---

alternative)... " (p. 12). Peer observation, according to this communication, was suggested as a source of information that can be valuable when made by at least "two faculty observers for each course." (p. 12)

When conducting the evaluation of teaching, departments relied on different sources of evidence<sup>2</sup> such as the instructor's peers, current and former students, and external observers. Information from students was collected through the campus evaluation system as well as student performance in the course. Some departments administered other surveys to students in addition to those used by the campus evaluation system.

The emphasis that the college legislation had on peer evaluation was reflected on the practice of evaluating teaching. Peer evaluation was given high importance by the College and its Departments. As a Department Chair, mentioned: "Peer review is important, especially if it involves observing several classes, because you learn more about how the instructor interacts with students."

Peer evaluations were said to be especially valuable for understanding the problems experienced by instructors who obtained low ratings. As the Department Chair added: "Most faculty members, at least 2/3 of them, are ranked below average during their first three years. Peer evaluation is important to understand why this happens."

When instructors obtained lower ratings, the department administrators had the responsibility to decide what to do. For example, a Department Head mentioned that when instructors obtained low ratings or had problems improving them, he talked to them as well as to the professor's students and colleagues:

When someone has trouble with his/ her scores, I talk to them. Sometimes the problem is not the teaching, but students perception that the instructor is arrogant because s/he has a formal way to approach them. Over time,

---

<sup>2</sup> According to one of the College's Associate Deans, not all departments used all the different sources of information for obtaining information about teaching quality.

most instructors become more informal in their interaction with students. This helps their ratings to improve.

Professional development for instructional improvement was also provided at the departmental level. Some departments used mentoring to help instructors in improving their teaching, but this did not take place in all departments.

Indeed, the college administrators had the perception that not all departments were dedicating efforts for improving teaching. As one of the College's Associate Deans mentioned:

Teaching mentoring takes place in some departments. Unfortunately, there are not too many departments that will provide it. In some departments, efforts for improving teaching move faster than in others. Some departments do it with rigor and care while others don't pay too much attention to teaching whatsoever. Sometimes you also find that some subsets of a department trying to improve while others aren't doing too much about teaching. An example of a department working hard in improving teaching through creative writing and mentoring is the English department.

To help professors to improve, departments were implementing different strategies based on to their own needs. They also relied on support for teacher evaluation and development provided through the College's Teaching Academy and other campus units described in chapter five.

### The Teaching Academy

In 1999, the College developed a Teaching Academy to provide peer evaluation and professional development to their junior faculty and teaching assistants. The Academy provided a nine-month enrichment program for new professors that included seminars, workshops, and informal discussions about teaching. According to the instructional specialist from the Office of Instructional Support and Improvement (OISI) who worked with the

Academy, all professional development provided by the Academy stressed the important relationship between teaching and learning:

The work of the Academy is based on the belief that there is a scholarship for teaching and learning. We work with faculty helping them in developing and stating the learning objectives, i.e., what students should know, understand, and be able to do as a result of participating in your course and classes, and then providing the instruction to ensure that this learning occurs, and then assessing whether and to what extent that learning has occurred. Curriculum, instruction and assessment are congruent in this model. The emphasis of the instructor is on facilitating the success of the learners.

The professors benefiting from the services provided by the Teaching Academy were mostly junior faculty members, but the Academy also addressed the professional development needs of other professors. As the head of the Academy stated:

We work with two kinds of faculty brand new faculty, (just hired), and with assistant professors and associates who have some minimal teaching experience. We try to motivate them and give them an opportunity of learning new ideas for improving their teaching and for improving student learning.

The Teaching Academy also provided faculty with a 24-hour retreat where professors were introduced to innovative instructional techniques, and follow-up workshops in teaching. During the workshops, there were sessions on student perspectives on teaching, balancing teaching and research, interpreting TES evaluation ratings, active learning strategies and the use of technology in teaching. In addition, Academy staff provided additional feedback for tenure. As an instructional specialist added:

The Academy also provides instructors with information about the how to achieve tenure, and about how to use the TES evaluations for improvement.

The orientation of the professional development provided by the Academy was consistent with the work of the Office of Instructional Support and Improvement since their specialists worked with the different Teaching Academies on Campus. The role of this office is described by the specialist who worked directly with the College of Arts and Sciences at Midwestern:

What we do here (at OISI) is focused on doing things in the best way for each of our clients. At the Division of Teaching Development, we all freely share the processes and products we develop and implement, so that our colleagues can benefit from our successes and learn to avoid mistakes we have made.

In addition to its professional development activities, the Teaching Academy organized a mentoring program. In this program junior faculty members were paired with senior professors known for the quality of their teaching. Some mentors reviewed TES evaluation results and discussed them with the instructor. As one instructor, mentioned:

The Teaching Academy gives us the opportunity of having a colleague observing our teaching. Mentors are selected among the best teachers of the college. Each professor is paired with one of these mentors. They observe our classes and provide us with feedback. Some of them even read with us the TES evaluations, and encourage us to have conversations with other colleagues for teaching improvement.

The Academy also provided junior professors with fellowships as support for instructional improvement projects, such as for development of instructional materials or techniques or on the application of technology in teaching.

The Teaching Academy conducted evaluations of all the different services provided to the faculty. Results from the evaluation were used to obtain feedback for

improving future professional development activities. As the instructional specialist responsible for the evaluation commented:

Last year, we conducted an evaluation of the work of the Academy with formative purposes. We tried to get the perspectives of different people from faculty who participated in the Academy to professional developers. After we collected the data, we looked for patterns, confirming and disconfirming what the developers thought. Then, we used the information to improve the services that we provide.

The Head of the Teaching Academy added:

The overall responses to our professional development were very positive. Participants suggested for us to include more interactive activities and to provide more opportunities for discussion. We also evaluated the mentoring program and the feedback was very positive. We are using the information from the evaluation to improve all the professional development activities that we provide.

In addition to the support provided by the Teaching Academy, professors at the College of Arts and Sciences had other instructional resources available, including the Office of Instructional Support and Improvement (OISI), the Midwestern Center for Writing Studies, and the Teaching Advancement Board described in chapter five.

#### Evaluation and Professional Development at the Departmental Level

The History Department at Midwestern was a department with a well-known reputation for the quality of its programs. In polls, it was nationally ranked among the best History programs in the United States. The Department was also well known for the number of teaching awards won by its faculty. According to the Department Chair, they had won more teaching awards than any other department on campus. One instructor commented:

History instructors are among those who receive more awards for teaching. This is not surprising because in order to be a good history teacher you

have to be good at story telling. Students like that. This is what makes you popular.

Department care for teaching was also reflected in discussions about teaching that took place across the Department. Another instructor mentioned:

The Department encourages good teaching. There is a lot of talk in the Department about who are good teachers. There are a number of very successful master teachers in the Department. These are professors who have received national awards and campus awards for their teaching. As young faculty members, we are reminded of this.

Departmental expectations about teaching quality were high, as the Department Chair put it:

In our Department all professors going for tenure need to provide evidence of good teaching. We enforce fair evaluation of teaching and demand a high level of teaching performance from our faculty members. We also have expectations for [superior] content.

Additional evidence of Department support for teaching was the work of two faculty members to help others teaching large classes. These two instructors wrote up a series of considerations for teaching courses with class sizes of more than 100. The considerations were adopted by the Department and were provided to all faculty members teaching these courses for the first time. One of the authors said:

With a colleague, I wrote a document that includes considerations for teaching large classes. It is now used by our Department but could be useful for all new faculty members at this University. We decided to write this document because when we began to teach, we realized that there were some important things for teaching large survey classes that most of us learned by trial and error. For example, one of the things that you don't know in advance is how to use the auditorium. You also need to know about all the kinds of students you get in these courses, and about rules and regulations. My colleague and I decided that there was a need for faculty to learn about this, so we wrote the document. Now our Department uses it for new faculty members. They can consult it and save time when teaching those courses.

The Department Administration had created a Teaching Awards Committee to reward good teaching. The Department Chair said:

We also have a standing Teaching Awards Committee that recognizes outstanding teaching. The committee is integrated, 2/3 students and 1/3 faculty members. This committee makes recommendations to the Department for awards on teaching. They also make nominations for College and University awards.

In addition to its reputation for good teaching and scholarship, the History Department was known for its progressive views. This was one of the first departments to support the unionization of graduate assistants. It opposed the use of Native American symbols as mascots in collegiate sports. As the Department Chair mentioned:

Our Department is very open to diversity. It has a very liberal orientation. For example, we did express our rejection of the use of Native American symbols as university mascots ... You can say that it is a Left Wing department. It is more liberal than others are.

#### Departmental Philosophy of Teaching

The Department was well known for scholarly orientation and for encouraging students to broaden their theoretical perspectives. A strong effort was put into a cross-discipline, cultural approach.

Although there was no written teaching philosophy, the Department Chair and other faculty members interviewed shared certain beliefs about the teaching of History. These beliefs reflected the Department's scholarly orientation and the need for teaching to be sensitive to issues of diversity. This was illustrated in the Chair's comments:

We want people who have an appreciation for complexity of historical experience and change, complexity of people's experiences, and the changes in the society. We want students to understand historical change too. Department objectives run in that direction, instead of expecting students

to master a particular body of knowledge. Diversity is very important for our Department, we want students to learn about the diversity in people's experiences through History.

A member of the executive committee who later became the new Chair, added:

We want students to learn about a non-Western culture and be able to write a research paper. We hope that students will learn about research methodology too, but I think what we really do is to teach them how to state a thesis, how to make arguments about historic events and periods of time. This is there, it is not written. We all want to get to methodology.

The approach to teaching History at the Department was consistent with national moves away from a universal view of history. The chair said:

There are two main orientations at work in the teaching of History as a discipline. The first one, which has great acceptance in the United States, moves away from universal theories when teaching History. For example, less emphasis is now put into comparing societies. The new approach is the study of different kinds of diverse experiences of people through History. This is the orientation that we follow at our Department.

The second orientation is a global, more comparative, international kind of approach, even when teaching US History. Under this view, we look at History as part of a society that we compare with other societies.

According to the chair, supporting one or the other departmental orientations had implications for teaching:

There are trade-offs when choosing one approach over the other. The acknowledgement of diversity is more honest in a sense. It is very difficult to make generalizations in History and it is very important for undergraduates to understand these complexities.

On the other hand, when you choose a more particularistic and cultural approach you lose the global, general view of History. By adopting a more complex view of history, we also create a problem for some of our graduates because, at the State [Education] level, they don't want them looking for complicated answers when teaching students about history. They want the right, simple answer that could explain a complicated event in History. This is a tension in the discipline too, on one side we are trying

to convey the complexity and the diversity of people's experiences while trying to teach students how to understand the change that takes place in a society and across history.

### Defining Good Teaching of History

In addition to Department use of a scholarly orientation to teach History, reflecting experiences of diverse people, the faculty and administrators valued certain characteristics in the teaching. Some professors shared the view that good instructors are able to adapt their teaching to different contextual conditions. As the department chair mentioned:

People who are willing to experiment are usually good at teaching History. Good History teachers constantly experiment. They are also student-oriented. They adjust their teaching to student needs and learning.

One of the award winning teachers of the college, added:

The best teachers of History are those who better adapt to their failures.

Good teachers were expected to adapt to the demands of different kinds of courses. One instructor commented:

You need to adapt your teaching to the kind of course that you teach and develop different sets of skills. Different courses require you to use different skills and conditions for teaching. If you are teaching, lecturing 400 students and you are at a podium, you do all the talking. In a seminar the role of the instructor is to run the class, to orchestrate an interactive discussion. The skills you need are different depending on the kind of class. In our Department it is common to know people who are good lecturers but not so good instructors in a seminar, and vice versa.

Good teachers were expected to adapt their teaching. One instructor said:

As an instructor, you have to adjust your teaching to different contexts and institutions. For example, I did very well at another university before I came here, but when I arrived I was a very bad teacher. I had to adjust to new expectations and student needs.

Another instructor who had received several teaching awards said:

I was at Washington University before I came to work here. I was getting awards for my teaching, but when I came here I did not do too well. The first semesters, I was a lousy teacher. The problem is that I was adjusting my teaching to other expectations. Students at Washington were different, so I had to adjust. Audience matters, if you are going to be a good teacher. With some audiences, I don't even know how to adapt. For example, I have a lecture-discussion class that always works well, and another that almost always never works. There is a senior seminar in the department that never works well for me when I teach it, students always saying that I am too hard.

In addition, good instructors were expected to adapt to the characteristics and needs of their students. One professor who had received national awards for his teaching said:

It is important to adapt your course to student diversity and experiences. You have to remember that different students learn in different ways. I try to be sensitive to this as much as I can.

The way in which instructors adapted their teaching was said to be consistent with the instructor's own teaching philosophy and style. One instructor said:

When teaching a course, many styles are effective. You have to work at finding the style that is natural for you. Some professors read class materials to their students while others use notes. There are some instructors who walk around the stage, but all can be effective or not. It depends on the person.

An award winning instructor said:

There is no one way to teach. Some teachers are up front, outspoken. Others are laid back, seated at the desk with students; other are rigid. The key is to find the way it works for you. Personally, I am good at teaching more for rigor and argument. I focus more on public policy issues.

Another award winning professor added:

I think it is important to be true to yourself, to know what you are trying to do and work towards your goals. There is not one kind of effective teaching. What works for one doesn't work for another.

A professor who later became the Department Chair said that instructors needed to realize that even when making changes for the better, their teaching was not going to work well with all students:

Not all instructors are good for all audiences. Even one of our instructors who had won national and local awards for teaching and who is an excellent instructor, even he has problems teaching large audiences. He is an extraordinary teacher. But even he can't have all his students liking his course.

Other considerations of good teaching. Members of the Executive Committee and other faculty members mentioned different aspects that they believed to be shared within the Department about teaching:

1. Good teaching promotes good citizenship. An award winning instructor stated:

At Midwestern University we are teaching citizenship; we want students to recognize different points of view, to know the diversity of experiences that take place and to value them.

2. Good teaching requires the instructors to be respectful of their students. The Department Chair stated:

Teachers have to be respectful to students. This is important because it influences student views of teaching and learning. Teachers have to be respectful when checking for understanding.

Another instructor added:

It is important to respect student ideas and opinions. I never say: "That is really dumb." Instead, I might say: "Ok, Tom says that he thinks George Washington... Do any of you have a different perception on this? Do the rest of you agree with this?"

3. Good teaching involves helping students to make connections with the content. As one award winning teacher mentioned,

I believe teachers have to be responsive to two things: (a) teaching certain material; (b) and teaching to people. No matter how much mastery you have of the materials or how many ways you use to teach it, you also have to work on how to connect this with students. This involves our attitude towards our students, telling them and making them feel that you are interested in making the course interesting for them.

Another award-winning instructor mentioned that helping students to make connections with the content could be done in different ways:

Undergraduate students need a context, such as a theme, a story, etc. so they can connect. There are a lot of people in our Department who are very good at doing this and there are others who have problems presenting the material in a way in which students can grasp it.

Another professor described how he was able to help students make connections with the content:

What works for me is to present everything that I do in terms of an argument. Sometimes, at expense of some simplification. For example: this is what I am arguing about. This is the evidence that I have to support it ... What do you believe about this? I do this in lectures. In seminars you use more argumentation and encourage students not to bring opinions instead. I challenge students in a way that, if they did not trust me, if they didn't know I am not trying to humiliate them, it would be a disaster. I pick up on something they say. I try to use teaching tricks and adapt them. I use them to focus the discussion, even at the expense of leaving something behind.

4. Good teaching draws on student curiosity. The Department Chair stated:

Good teaching involves sparking a sense of curiosity in the student not only in that course but in the learning process, a curiosity that leads him or her to break through what people said in class and reflect on it.

One of the national award winning professors added:

There are few students who will become historians. If students know the date of a historical event, it is OK, but what is important are the tools they have to have for issues and questions in the materials. It is important to

figure out where they are ... to say where you want them to be. Then, you have to figure out how to help them move there.

Another award winning instructor said:

As an instructor, I have been able to cut through the alienation of students. Sometimes, I have students who don't want to be there, who would prefer that I teach as if I was the textbook, but this is not the way to teach. The key for me is how to get them where I want them to be. There are people who are very good at having students liking the course. But this may not be good enough if you know they are not learning what you want them to learn and with the depth that you expect them to learn.

### Evaluation and Professional Development at the Department Level

The evaluation of teaching in the History Department followed campus and college policies using annual review, tenure and promotion procedures. The Associate Dean of the College said:

During annual and third year and tenure reviews, departments look at teaching. The Executive Committee of our Department takes a long look at each evaluation source. We pay attention to the context of teaching.

Evaluation of teaching required the use of different sources of evidence, such as student ratings (TES ratings and a Departmental survey of students), peer observations, and document review. According to the Chair, the Department looked at different sources because each source provided a different kind of information:

We use student ratings and peer review with other contextual information when we make decisions about teaching. We have to use both peers and students because they bring different views. Student ratings are important because you need to get a sense of how students feel. Many professors are skeptical about them, especially when using rating results for making tenure decisions. They feel that the ratings are not always appropriate for measuring instructional quality. Peer evaluations give you a different perspective, the perspective of an experienced teacher.

The Department Chair also stated that the use of a single evaluation source, such as student ratings of teaching could be problematic:

Most faculty members will reject using students as a single source for evaluating their teaching because of the required nature of some courses. In some cases students may not be the best judges. That is why you have to have experienced teachers observing someone's teaching.

The members of the Executive Committee identified different benefits and limitations of student ratings and peer evaluations, as well as other sources of evidence for evaluating teaching.

### Student Ratings of Instruction

The Department obtained information from students through the campus main evaluation system (TES) and through a survey administered at the end of each semester (see the appendix for a copy of the forms). TES ratings were required for junior faculty. For promotion, tenured professors were not required to provide TES ratings to document the quality of their teaching. One of the College's Associate Deans said:

By obligation, all junior faculty members need to use TES evaluation forms, at least for two semesters. TES forms are required during the tenure process. Professors going from associate to full professor are not required to use TES evaluations. The TES is optional for tenured faculty.

The Department Chair added:

All untenured faculty members are required to administer TES evaluation forms. Results from the TES evaluations are reported to the Department. Professors going for tenure and for promotion are encouraged to administer the TES surveys. We use the scores on an annual basis. All untenured instructors have to administer TES surveys as well as collect evidence of research productivity. Results of this review are looked at by the Chair and reviewed by the Executive Committee who meet with the candidate. During the meeting part of the discussion focuses on the evaluation of teaching, and about how teaching might be improved. The Department

posts the [campus] list of instructors rated excellent by their students. This is a way to acknowledge the instructors' efforts. When tenure decisions are made, the ratings are also included as part of the teaching record.

According to the Department Chair the reviewers differ in the weight they give to the campus evaluation forms:

When a professor goes for tenure, three professors look at the instructor's records. Each person brings something different to the evaluation and gives more or less importance to some sources than others. I take the TES evaluations very seriously, more than other members of the committee do. Most committee members give less weight to student ratings because of the problems associated with the ratings.

Usefulness of student ratings. The Department Administrators had identified some benefits from using student ratings. As a member of the Executive Committee stated:

TES ratings provide more useful information about the General Education courses than about those courses in the Major. Some professors don't think they are really useful. I personally believe the ratings aren't useless when they are read well.

The ratings were found to be useful as indicators of problems in a course. According to a member of the Executive Committee, student comments were especially valuable for understanding the scores:

The numbers are an indication that something I did went wrong. For example, I had a semester in which I got lower TES rating scores. Later I realized that this was because of the way in which I approached students. There were a couple of student comments that made me realize that this was what happened. So when I was looking at the low scores in the report, the comments made me realize what went wrong.

This perception was shared by one of the award-winning teachers:

The comments made by students in the evaluation forms are useful and helpful. I ask people to write and give me feedback. Then, I look at their

comments about what was good or not, about what could be better. The number is the less useful part.

Another member of the Executive Committee claimed that student ratings and peer observations provided similar information about teaching:

There is a raw correspondence between the ratings and peer reviews. If someone receives low TES scores, I can tell after a few minutes of observing his/her class why this happens. The instructor may have included very good classroom materials, but if s/he is too harsh to students or has problems connecting the lesson with other aspects of the course, then you understand why s/he gets low scores. Among the faculty members that I have observed, when I see high TES evaluations with a couple of exceptions, the teaching reflects the goodness.

An instructor, however, stated that TES scores were not useful for evaluating some courses:

TES scores work better with some courses than others. They are fair, fairer for small classes. In general, students talk about what they like or dislike the most of a course. If they don't like my text, I may use it in a different way, and this may work fine.

Problems when using student ratings. The members of the Executive Committee and others of the faculty had identified certain problems when using TES student ratings in making decisions about teaching quality. Some of the main problems were:

1. Students tended to rate high those professors who used more instructional technology. As

an instructor commented:

Students like technology and expect you to use it in the classroom. There is no problem for most faculty members to use audiovisual materials or some kind of technology but sometimes students just want to be entertained. They take your course because they need it to fulfill a requirement, not because they really want to learn history. So, when you receive the evaluations, it is not surprising to hear them saying that the class is boring, that the TA's are boring, that you are boring.

A member of the Executive Committee sharing this perception added:

Students like the use of technology. They like to be entertained. They are the MTV generation. They grew up spending a lot of time in front of the TV, so they want images that change all the time. Spending a night reading a book is boring for them. We make a strong effort to make the courses more interesting and to introduce more audiovisual material but this is not going to be the History Channel. Students have to do their work. They have to read and make interpretations of the readings. This is not *Survivor*. It is not just not entertainment.

The Chair, although supportive of the use of technology was cautious about it:

Our Department supports the use of technology, but relying too much on technology could lower the quality of education, especially if people use technology as a substitute for good teaching.

2. Student ratings were not always indicators of good teaching. As a member of the

Executive Committee mentioned:

TES results don't have meaning by themselves. Sometimes, I have observed someone who obtains very high ratings but when you see them teaching you don't agree with the ratings. Other times you get someone who is receiving mediocre ratings but you can see that the scores don't reflect the goodness of the teaching. So, in some cases the scores reflect good teaching. I have seen people who are terrific and who obtain high evaluations. But, there are also times when you observe the class and you can tell the instructor will get spectacular results in spite of the poor work that s/he is doing at teaching.

Another member of the Executive Committee illustrated how the TES ratings were not useful for indicating instructional quality and did not indicate that the instructor had failed to provide students with something needed for a later course:

I have in mind a professor who teaches in a certain way that helps him obtain good ratings. But when his students come to my class, they arrive with the idea that they just have to read the book in a certain way or write their papers in certain forms to get a good grade. In cases like this, the scores do not reflect that the course did not provide students with something they needed. I have to do a lot of undoing in these cases.

The use of averages of the scores was said to be the least meaningful part of the TES evaluation reports. As an instructor mentioned:

Averages--low or high--don't mean anything. You can't do anything with them. You need to know more about the teaching.

3. Student ratings were seen to be influenced by gender, race and age and were said to be especially harder on young woman. The Department Chair said:

Over the last 10 years, we have found that the evaluation [by students] can be influenced by two main variables: gender and age of the instructor. It is especially hard for young women. Students carry preconceptions when they come to the classroom. I am almost positive that it affects the assessment they make of the instructor's teaching. I believe the assessment is hard for women and even harder for young females. If you have taught for a lot of time, you get used to this.

A female instructor, also a member of the Executive Committee, stated:

Equity is an important issue to which our Department is sensitive. Students are hostile to women and non-white faculty. . . . I think racism is an issue for the evaluation.

Another female instructor added:

Some students are resentful of women in a position of authority. This creates serious barriers for teaching. It also works the other way around. For example, Latinos and African American students may work well with an instructor who teaches a course that is related to their race or ethnicity because of their genuine interest in the course.

A minority instructor shared this perception:

Gender and age are an issue in the evaluation of teaching. You always get comments that reflect students prejudices. Students make these comments even if they don't understand the comments they are making. The classrooms are not isolated from the society; they are a reflection of the American society. Students bring assumptions on the behavior of others, of who is smart and who isn't. It is always there. IT is something you have to respect and deal with it.

Another member of the Executive Committee added:

I remember a teacher that I had when I was a student. She was young and came to class in a mini-skirt. This was Harvard and you had these jerks laughing at her . . . You expect a woman to make a mistake . . . This happens here too. It is not surprising that women get lower scores than men in the TES evaluations.

4. The Department Administrators also found that student ratings were influenced by the instructor's rapport with his or her student. A member of the Executive Committee said:

People who can connect with students can obtain high ratings even though they aren't good at teaching. I have also observed TA's who receive very high ratings but you see their teaching and realize that their content coverage is very superficial. They have good rapport with students but they are poor teachers. That is why you need to look at the teaching.

5. In addition, the Chair noted that grade expectations also could influence the ratings:

A few years ago, our Department did a study because of the steadily rising number of students transferring to our Department from other majors. (A total of 24 % of those transferring were either on probation or had been on probation early in their careers.) Findings of the study indicted that there was a significant spread of "A" marks awarded by our teaching assistants who were in charge of the General Education courses. We also found a correlation between having high expectations of grades and student decisions of transferring to the department. After conducting observations of some TA's teaching we concluded that our TA's needed mentoring in both teaching and grading. We created an ad-hoc committee that could supervise large number of our TA's and provide them with guidance.

6. Having an accent different than that of the students was also identified as a factor that could negatively influence the ratings. A member of the Executive Committee said:

Students react to your accent ... I have a Western accent ... I am aware that they react to that. It can affect your ratings. But there are things that work if you present the material in certain ways. Negative perceptions aren't immovable things, you can use them to your advantage.

7. The Department also had found that class level and size could negatively influence the scores. One of the members of the Executive Committee said:

When looking at the ratings, we have to consider class level. It is easier to get lower ratings in a 100 level class than in a graduate seminar. Class level matters.

Most 100-level courses had a class size of more than 100 students. Class level was related to class size. Another member of the Committee said:

The size of the class is important when making decisions about an instructor's teaching quality. History courses vary in size. Some courses have 300 - 400 students. Other courses have 35 students that allow more instructional interaction. Seminars have approximately 12 people. We also have introductory courses with 1000 students divided into six groups. Class size influences your ratings.

The Associate Dean of the College added:

It is common to obtain a score of 4.5 or 4.6 in a graduate class with 5 - 15 students. Anybody could get this score in a class where there are only a few students. Small classes usually obtain high ratings. Generally, in a small class, it is almost impossible to get bad ratings.

Another instructor added:

Large classes are especially difficult to teach because it is very hard to have student interaction. In large classes, (600 students) I don't do well. Most instructors don't do well in these large courses either.

8. Committee members also stated that sometimes students were not in the best position to make judgments about certain aspects of teaching at the time of the evaluation. A

Committee member commented:

The main problem is that students don't always know what is best for them. Many of them prefer less reading over more reading, even though that is not good educational practice. I think students should have some

academic pressure. When students say, "too much reading", you need to look carefully at the teaching. This might or might not be the case. Students also have different ideas of what material they need to know. Ten years later they may understand why some curriculum choices were made and why some reading materials were selected.

9. The way in which the TES scores were reported was also identified as a problem of the campus evaluation system. As the new Department Head stated:

One problem with student ratings is the forms used for reporting. I have never seen a more illegible student evaluation form. The TES forms are particularly cumbersome.

10. Other problem of TES ratings, according to the members of the Executive Committee, was the lack of agreement among student perceptions of "good teaching." Some of them thought students ought to agree:

One of the problems with a mass evaluation of teaching is that different students have different expectations of the instructor. They see "goodness" in a different way. For example, I was present when the head of our college's Teaching Academy asked undergraduates about what they thought was good teaching, about what they wanted or expected from a good instructor. I remember one student saying that for him good teaching meant that the instructor delivered the material in an organized and clear manner. Another student said that a good instructor had to be prepared and on time. Other student said that the presentation has to be exiting, fun, and dynamic. A forth student said that for him, teaching was only good if it inspired him. A fifth student indicated that teaching was good only if it made a difference in his life. What worked for one didn't necessarily work for the others. This is why it is hard to measure effectiveness. Mass evaluations of teaching have this problem.

11. One of the college award-winning instructors recognized that a problem with the ratings was that they were not necessarily linked to what students were learning:

I don't put too much weight on the TES ratings, even though I get very high ratings. It is a statistical measure, but it doesn't tell me how well my students have learned. I am a little skeptical about the scores. In my evaluations, some students complain that I ask too much from them, but I

feel that what students need to learn is important. I am firm in my position of keeping what is important for them. We are not car salesmen; we are not here for entertainment. A lot of people confuse that and that is not good.

Some instructors obtaining higher ratings had a more positive view of the ratings. The

Department Chair mentioned:

Criticisms of student ratings are sometimes true but weaker teachers --the ones getting low scores, are the ones complaining about the ratings. Those who do better are more sympathetic.

The Department administered, in addition to the TES student ratings, another survey to its students (see appendix C). One of the members of the Executive Committee said:

We have a separate form that we administer for History classes. This evaluation is independent of the TES forms. I find this form to be more useful for me because it asks students to comment on the course reading. It is more valuable than the numbers that you get from the TES reports.

Other professors, however, found these surveys to be not very useful. They had students make a judgment about the course readings, and the professors did not believe students were in the best position to make this kind of judgment.

### Peer Evaluation

In addition to collecting information from the professor's students, the Department had created a peer review process. The Department Chair provided this brief overview:

Our History Department has a qualitative procedure for evaluating teaching. It has been in place 10 years. We have a peer evaluation system in which members of the Executive Committee are selected to observe the instructor's teaching. The selected members observe the class and write a report that becomes part of the instructor's teaching records. A copy is handed to the instructor being evaluated.

Peer evaluations of untenured faculty take place at least once each semester. For tenured faculty, peer evaluations take place when the instructor is going

for promotion. We rotate the members of the Committee who do the observations, so different people evaluate the instructor.

After the evaluation, peer mentoring takes place but it is not very systematic. There is not a structure to train faculty in the Department. What we do is to have young faculty members talk to other professors (young and old). They share ideas of how to present information or about content coverage. There is a culture of sharing in the Department. It is a culture that is open to different ways of teaching.

The kinds of information included in the peer reports varied depending on who conducted the review. For example, the Chair stressed classroom organization and student involvement:

Whoever is doing the report pays attention to the characteristics of the class (lecture, organization, clarity of delivery, etc.)--when the class is a lecture. If a discussion, we pay attention to how the instructor involves students in the discussion. We look at class materials such as overhead transparencies, sound, etc.

Another Committee member said to be paying more attention to student response:

I think what we pay more attention to is instructor effectiveness: Are students taking notes? Are they asking questions? Is the lecture too sophisticated for them? Is it too boring? Is it too detailed? Is it non-argumentative?

Another member of the Executive Committee focused more on scholarly aspects of the teaching:

I have evaluated people's lectures and coordinated people who teach. When observing someone you have to ask yourself: Does it reflect the state of the field? Does it show that you are a good historian, not only a teacher? Is it engaging? Are you teaching not just facts but finding ways to teach methodology and thematic approaches to the materials? Are students thinking critically, writing well, and knowing how to research a question beyond the surface of issues?

These differences in definition of good teaching did not seem to be a problem for the Committee making decisions about teaching quality, it integrated all the information:

Benefits of peer review. Having a peer observing one's class and providing feedback<sup>3</sup> was said to be very useful for instructional improvement. As the Department Chair mentioned:

We conduct classroom observations of teaching and use anecdotes from the observations in our teaching report. We think this is better for instructional improvement because the observer can make recommendations to help the instructor improve. Some peer observations tend to be only positive. The good ones point out the weaknesses and strengths of the instructor.

Observing a peer teaching a class was said to be also useful for the observers. A member of the Executive Committee, said:

In general, for all teachers it is good to look at others teaching. In my case, my wife teaches and that is very helpful. She is a better teacher than I am. One time I taught with her, it was a great teaching experience to look at the way in which she presented the material. Observations are different when you are a student.

Peer evaluation results were found to be valuable for interpreting student ratings results. One member of the Executive Committee said:

Peer evaluations help us to pay attention to specifics. They help explain the TES ratings. Peer reports bring the context. They bring a narrative of the teaching that TES don't. TES scores don't tell you about the level of difficulty of the materials and the integrity in which the class is taught. In other words, you don't know if the instructor taught the complexities of the subject or just taught a happy story with a lot of laughs ... When you use peer evaluation you know if the professor talks at a level appropriate for the students.

---

<sup>3</sup> The Department also conducted observations and provided feedback to the graduate teaching assistants in the Department.

Another Committee Member claimed that peer reports provided a better reflection of the teaching than the TES ratings:

As a Committee we can not ignore the TES ratings because they are required. So, we look at them first. We do pay attention to them but we give more importance to the peer reports because peer reviews are thoughtful reflections on someone's teaching. We give less weight to the ratings because TES ratings can be deceiving. For example, if you get low TES scores it can be that a lot of things went wrong. It could be that a lot of people liked your class but just a few hated it

According to the Department Head, peer evaluations were useful for learning about the ways in which the instructor treated his/her students:

We pay strong attention during observations to how the professor's respects his or her students because this influences student views about the course. We also pay attention to the kind of courses taught by the instructor. The observations provide evidence that someone is very weak in large lectures but good at advanced seminars. We can know all of this if we observe various courses taught by the same instructor.

Peer evaluations usually supported the findings from the TES ratings. According to an instructor:

Peer evaluations and TES ratings aren't all that contradictory. People who get low TES don't obtain glowing narratives of their teaching. There are exceptions. For example, there are professors who get high TES ratings but when you observe them you realize that they don't teach well. There are also those who don't get high ratings, but when you observe them you realize they are doing a good job.

Problems of peer review. Members of the Executive Committee also found some problems associated with the use of peer evaluations of teaching. One of the problems was the conflict of interest faced by the professors who observed the instructor. Because all the observers were members of the Executive Committee, they felt compelled to provide positive evaluations. A member of the Committee acknowledged:

Professors find it difficult to be critical of one another. Observation reports are usually positive. You just don't get a professor telling you that there are a lot of problems with the teaching. This just doesn't happen. It is part of being a colleague. As a fellow professor you are guarded about making negative statements. It is a normal human interaction. Sometimes, however, you can not avoid mentioning negative aspects affecting the quality of student learning. In that case, we just include them in the report.

The Department Chair added:

The Executive Committee selects who will do the observations. So, when *you* sit in the classroom *you* also have a stake. You are a member of the Executive Committee and the Committee is the one that recruited new faculty. If you recruited a person, you go and do the evaluation with that in mind. If someone doesn't get tenure, it is destructive to the Department, especially if it is a department that you are trying to build. As a member of the Executive Committee you have an interest in keeping the instructor. You want them to do well.

Another problem of peer evaluation according to one identified by a Committee member was that of identifying teaching quality:

You need to pay attention to whom is writing the peer observation reports. Many peers are not critical about the teaching. Many of them just say something nice about the instruction. There are also people who can't recognize good teaching when they see it.

In addition, the Executive Committee found that peer observations could be influenced by student ratings. Even though observers did not see the scores, they knew about instructor reputations as a result of comments made by his or her students. One member of the Committee said:

I don't see the TES scores when I go to observe a peer teaching a class, but I usually know something about the scores prior to the observation. Sometimes they already had told me they had a problem with the scores. Other times, you know something about the TES ratings because they are in the list of excellent teachers.

Another Committee member believed the effect of knowing about the TES evaluations was minor. She said:

When doing the classroom observation you haven't seen the instructor's ratings, but you know the instructor's reputation. This probably has some effect on my evaluation but I think it is a minor effect.

### Other Sources of Information Used to Evaluate Teaching

In addition to student ratings and peer review, the Department paid attention to student outcomes, classroom materials and artifacts. They also looked at the instructor's examinations. A committee member said:

When we evaluate teaching we also look at the teaching statement to see the instructor's approach and philosophy. We look at his examinations to see if the instructor is being effective in assessing his course. We look at the syllabus, the list of readings, etc.

Some professors also looked at student work as information about teaching quality. An award-winning instructor said:

I pay attention to the quality of student work when evaluating teaching. I even do that with my own students. I use it as feedback about my teaching. By looking at how students did on classroom assignments, you learn how to help them do better.

Another instructor added:

Looking at student papers helps us understand in what ways the instructor is being effective. Personally, I change my teaching as I learn more about student learning. But I don't sacrifice covering things that matter in their lives.

According to the Department Chair all the information collected from all the sources was expected to be included in a teaching portfolio. All junior professors at the department were encouraged to develop these portfolios, with help from their mentors:

Individual professors are expected to prepare teaching portfolios describing their teaching philosophies, and to include evidence of their teaching, TES ratings scores, classroom materials, etc. They are supposed to keep up the portfolios, so we can see improvement of teaching over time.

A member of the Executive Committee added:

All professors going for tenure and promotion are expected to have portfolios. As Committee members we are encouraged to look at the evidence of teaching included in the portfolios, such as lecture outlines, tests, student ratings, peer evaluation reports, etc.

Teaching portfolios were updated during the third year review and later during tenure and promotion cases. One of the Associate Deans of the college of Arts and Sciences commented:

In the History Department they don't rely only on TES ratings. Every semester they have a peer observing the class of non-tenured faculty members. They also collect documentary information about the teaching. All the information collected from the evaluation is part of the teaching portfolios. The Department revises the teaching portfolios of all non-tenured faculty members during their third year review. The review is useful for instructional improvement. We have found that professors become more conscious about teaching by keeping the portfolios. The process of preparing the portfolios helps them.

These administrators said they believed the departmental evaluation system had limitations and strengths. An Executive Committee member said:

Even though the Department evaluation system presents some problems, it is a wonderful and well-structured system that provides you with an internal evaluation of teaching and information from peers.

The Chair added:

The system works well. Sometimes senior faculty members tell us that we give too much support but we think it important. We intend to have high standards and bring in people who can perform at that level. But, we also support them so they can do well.

## Making Decisions About Teaching Quality

The evaluation of teaching at the Department was conducted for two quite different ends, making personnel decisions and for improvement of instruction:

The evaluation of teaching affects decisions of tenure and promotion, so it is more important for junior faculty. The information from the evaluation is also used for instructional improvement. We hope that professors will use the feedback for reflecting about their teaching and working with their mentors in identifying ways to improve (Member of the Executive Committee).

When making judgments about teaching quality, Committee members looked first at the TES ratings and analyzed other important contexts. The Associate Dean<sup>4</sup> said:

First, we look at the ratings and put them into context. For example, we look at the kind of course. Required courses usually get mediocre student rating results because of the negative attitude of students towards the course, so we pay attention to this and other contextual factors when reviewing the teaching portfolios.

In addition to the ratings, the Committee also looked to the peer evaluation reports and other information from the teaching portfolios. The Department Chair said:

As committee members, we share copies of the observations and the summaries of TES results provided by the instructor. We also have activity reports for each instructor.

When the instructor's TES ratings were low, the Committee met with the instructor to discuss what went wrong and made suggestions for teaching improvement, such as observing

---

<sup>4</sup> The Associate Dean was also a professor at the History Department and a member of the Executive Committee.

another instructor's teaching, attending teaching workshops, and so on. The Chair said:

If someone obtains low ratings 2.5 points or less on the TES forms we go and talk to this instructor at the end of our review. In general, we ask the instructors how they felt their teaching went that year, what they thought went wrong. We ask if they want to observe someone who is being successful at teaching, someone who has received teaching awards or who is in the [campus] list of excellent instructors, or we encourage them to take a teaching workshop. In this way, instructors can get ideas for improvement. However, for our faculty to have low ratings is very rare. Most professors in our Department get average ratings. The TES scores of our professors are usually above 3 points and close to 4 points (on a five-point scale).

According to the Department Chair and members of the Executive Committee, most instructors who obtained low evaluations for their teaching improved over time because of the Departmental emphasis on improvement. When a professor got low evaluations for more than one semester, he or she was expected to work with the Executive Committee to develop of a teaching improvement plan. A member of the Executive Committee said:

When an instructor obtains poor evaluations for his teaching, we schedule an improvement plan. We might suggest the instructor look at other teachers. At other times, the chair makes suggestions for improvement. Teaching is important, if you don't do well on teaching it can impair you from getting tenure.

The Department had a mentoring system to support junior faculty members during their tenuring process and to help them improve when obtaining poor teaching evaluations. Mentors were expected to provide advice and guidance on teaching, research and tenure. As part of the improvement plan, professors could also work with their mentors in identifying ways for improving their teaching. The Department Chair said:

Results of the evaluation are used for mentoring. We have a mentoring system in which all untenured faculty members are peered with someone in their field. Each professor who is going for tenure chooses the person who will work with him or her. Sometimes they choose one person of their

same ethnicity, culture or gender. Advice and support is given in teaching and other things, such as writing or getting funding and tenure. Sometimes the mentor sits in the classroom to observe the class.

Helping another professor to improve his or her teaching was not considered teaching.

It was seen a service the professor did for the Department. A member of the Executive Committee said:

Helping others to improve their teaching does not count as teaching. Some things count more than others ... We have the mentoring system but very little help for teaching exists in the Department. I can't even think of a case in which someone has helped another professor improve his teaching, at least not in a formal way.

The Department invited instructors to obtain mentoring from other campus offices.

One of the award-winning instructors who had used the services of the campus Office of Instructional Support and Improvement commented on the benefits.

OISI is very good, particularly the head of the Division of Instructional Development. She is terrific! I usually go to OISI for help. They have provided me with advice on how to state my course goals and on how to rewrite them in terms of students' perspective. It was very useful. I use the services of OISI a lot. They are a worthwhile organization

The Department, however, did not have a policy of using the services of OISI. It did not send its faculty to receive instructional support. The Chair said:

I don't remember sending anybody to the Office of Instructional Support and Improvement. If we decide to ask OISI for help, we will need to tell them what we need. So far, we don't need to do this because we have our own mentoring system in place. I know that they videotape instructors and provide feedback for teaching improvement. I think that it can be useful for instructors to see themselves teaching. However, we don't use this resource because their emphasis on teaching, especially undergraduate teaching, is a prepackaged approach. We believe that teaching can vary from one period of time to another. If we decide to use the services provided by that Office, it would be customized within our Department's expectations and needs.

### Consequences of the Evaluation

According to the Department Chair and members of the Executive Committee, there were several rewards for instructors who obtained high evaluations of their teaching. One was to be included in the campus list of excellence in teaching. And professors could apply for teaching awards at the department, college and campus levels. For those instructors receiving poor teaching evaluations, there were other consequences. The most common was to have their reputation hurt. They might be denied tenure if their research was not above standard. A committee member explained:

Some people are so unsuccessful at teaching that this can be a problem. But, professors are only at risk of not getting tenure when their research is on the borderline. If their research production is good but not high, then, in this case poor teaching receives more attention. If the research is not so good and teaching is not very good either, then teaching could be one more thing against the instructor and this could result in denying tenure.

According to another member of the Executive Committee, there was little possibility for a weak-teaching instructor to be denied tenure:

Teaching isn't a major card in the tenure decisions. What usually happens when we have a very good scholar with lousy TES scores, we try to see how to help him/her to make his/her case. We use the peer review reports and see if there is some evidence of teaching goodness. We think, "How can we help this person whose scholarship is so good to get tenure, given that the teaching ratings are low?"

Thus at times, the reason for improving teaching was to preserve scholarly research. The Department Chair added:

It is common to find good scholars who aren't so good at teaching. But, you find that their scores are not really bad and that they are making efforts to

improve. This is the common situation. So, if someone is a good researcher, teaching won't be a problem. S/he will be fired only if there is evidence that students are at risk.

Consequently, given the research orientation of the university, the possibilities of being denied tenure because of poor teaching performance were seen as very improbable.

### Teaching Versus Research

Although the Department Chair and members of the Executive Committee stressed the importance of teaching, they recognized that research had a clear priority over teaching, especially apparent when making decisions about tenure, promotion or salary increase. The Department Chair said:

Teaching is important but it doesn't eclipse research. The University research mission is very important. Like other Departments, we pay attention to teaching when we assess merit every year. When we review the files for annual salary increases, we make an assessment of publications and teaching quality. Yes there is more emphasis or value put to research but we take teaching very seriously. Teaching is not the most important part of the decision but it is a significant part of it.

Another faculty member stated the importance of research over teaching:

Research is still more important than teaching when talking about tenure and promotion. Research is also more important for building a reputation outside the Department. The most important historians in the field are recognized because of their scholarship. Scholarship counts more for the professor's advancement. Teaching gives you another kind of gratification. For some, teaching is more gratifying.

In general, when a professor was going for annual review, tenure or promotion, the Executive Committee looked at all the work done by the professor, giving priority to scholarship. If the professor's research productivity was high, then teaching had a secondary role in the review. One professor said:

What really matters in terms of your pay and status is your scholarship. If you are good in research, they don't worry too much about teaching. This is a research university and research has priority. Should we tell people this when they come to teach? No, Some people may even consider this to be something negative.

One of the Associate Deans added:

Sometimes, someone will come up with very bad TES ratings in their review but s/he may be very good in research. It will be more feasible for this person to get tenure than for a person who is an excellent teacher and a poor researcher.

A member of the Executive Committee provided an illustration:

Last year we had to make a decision about an instructor going for promotion. She is a very good scholar, but her TES evaluations of teaching weren't very good. We were wondering what to do in this case, but she is an excellent scholar and we knew that she was trying hard to improve her teaching. We looked at her teaching statement, her teaching philosophy, and her examinations. In the end, we granted her tenure.

If research production is good, evidence of good teaching helped the instructor get tenure. Teaching performance acquired more importance in cases where research production was not very good. Another member of the Executive Committee said:

Research means more when making administrative decisions but it doesn't mean teaching doesn't mean anything. If you have someone who is doing well in research and teaching, you can also use evidence from teaching to support the tenure case. But, if you are an average researcher and your teaching is poor, poor teaching can make things even worse for you.

If you are a great researcher, you have to do something outrageous during teaching to be affected, such as attacking your students. But if research production is not high, then poor teaching can be what is needed for denying tenure.

Balancing teaching and research was considered difficult to accomplish. Two award-winning instructors were among those who believed that a professor could succeed in both.

One said:

It should not be a conflict between being a good researcher and being a good teacher. Teaching and research go together, they complement nicely. You can share the excitement of research with your students.

The other, ignoring views that research and teaching are fundamentally different, said:

There are faculty members who are dedicated scholars but poor teachers and vice versa. There is no reason you can't do both well. It takes a lot of time and it is very stressful but you can do it. You are a better teacher if you are a researcher with a research agenda, and you are better researcher if you teach.

The Chair acknowledged the fundamental difference but said they were not mutually exclusive:

Most professors do their writing during the Summer or during a sabbatical, but teaching and research are not mutually exclusive. Scholarship is something very private, as a Historian you can spend your time in the archives. Teaching is a way of serving the society, a way of going public with your knowledge.

These advocates of research cited no research on the compatibility of teaching and research. They did not quote distinguished teachers. They drew from personal experience and repeated the slogans often used to ward off a Legislature voicing concerns about teaching quality.

### Emerging Issues

Some issues identified at campus and college levels were also identified by the Department Chair and members of the Executive Committee, as important for the evaluation of teaching in the Department. These issues were:

Problems when measuring student and teaching outcomes. One of the issues identified by the Department was about external pressures for measuring outcomes, both student and teaching outcomes. A member of the Executive Committee said:

We are being forced to evaluate student outcomes. I am part of a committee in charge of this. We are pressured to measuring outcomes and we are working towards better ways to do this. So we came with the idea of comparing student papers. We kept some copies of student papers from their junior year and then compared them with other papers from their senior year.

The problem was that the papers were different. There were enough diverse expectations of what students should do when writing the paper. We found a lot of problems making comparisons. The papers were different in different aspects of what History is and what it means to be a historian.

Even more interesting was that the papers they wrote in their junior year were sometimes more interesting than those written in their senior years. It was a problem to use those papers for evaluation. We did not get anywhere. We did not learn much about outcomes by comparing the papers. At the end, we still have the pressure for determining outcomes and are still looking for standard ways of evaluating students at two different times. But another problem will remain. How much will any change we find be due to our teaching and not to the many [other] variables influencing learning?

Measuring teaching outcomes was especially problematic for some faculty members.

Another member of the Committee stated:

It is hard to measure "effectiveness." I have problems when people try to measure the goodness of an instructor by looking at student outcomes, especially because what makes a difference in the lives of students is not something that can be done in a hurry. Effective teaching is not well measured by looking at student outcomes. The problem with the evaluation of teaching outcomes is that you can't always wait ten years to see the benefits of the teaching over time. There are also new ways of evaluating teaching that can be criticized for being very subjective. I am not convinced that we can objectify good teaching.

Equity issues. Over a period of ten years, the members of the Executive Committee and the Department Chair found equity issues that deserved attention. They found that instructor race and ethnicity could play a role in the TES evaluations of teaching. The Chair commented:

Over the years, we have found that student evaluations of teaching are harder for “diverse” faculty. We have a problem with Latino faculty whose first language is not English. I remember a case when we hired a Latino professor and had some students complaining because of his accent. Anecdotal evidence exists that people who do not speak Midwestern English may get hurt in their ratings [at Midwestern University].

The Executive Committee also found that although instructor race and ethnicity could be related to poor teaching evaluations, this would not be an issue for all individual instructors. As the Chair described it:

Race and ethnicity are an issue in the TES evaluations but not in all cases. We have some African American instructors who do just fine in the evaluations but we also had cases in which students complained about an instructor and we believe that it was because of the instructor's race. For example, on one occasion, some students complained about an African American instructor because they thought he was biased when talking about American history. In that case, we did not pay attention to the complaints.

We had another African American instructor obtaining poor ratings but it was not a race issue. What happened was that he had very strict rules in the classroom and was hard in grading. He did not allow students to be late to class, etc. The interesting thing about this case is that we realized what the problem was when we observed the class. The students never brought up any problem in the evaluations. They just rated him very low.

Another member of the Executive Committee added that one of the problems with identifying student bias against instructors of certain race or ethnicity was harder when looking only at TES scores:

There is no doubt that race can influence TES evaluations in our Department but you mostly have anecdotal evidence of this. Unless students verbalize their biases, it is hard to prove that there is a bias. That is why it is so important to investigate when you have a claim of this nature.

The Department Administrators had found that having an accent could be a problem.

In some cases having an accent seemed to be related to low ratings but there were instructors who obtained high ratings in spite of, possibly because of, their accent. The Department Chair provided an example:

We know that students don't like to have instructors who don't speak Midwestern English. Having an accent can be a problem for an instructor. It can negatively influence your TES ratings. But, we also have people with an accent who do well. For example, we have an Asian instructor who speaks with an accent and gets great evaluations. This is a Chinese instructor who was born in Hong Kong. When we hired him to teach a course in Asian History, we were concerned about his thick accent. So we asked another instructor who was a native speaker of English to co-teach the course with him. The American instructor was very well organized, rigorous, impeccable at the delivery, with good overheads, etc. The Chinese instructor's language was intelligible, but he had a heavy accent, and was pretty disorganized and not very good at using teaching aids. At the end of the semester the Chinese instructor obtained higher ratings than the native speaker of English. Students were attracted to his enthusiasm for teaching and his knowledge of the content. I have seen this happen in other departments, such as Math and Chemistry where the instructor's passion for the topic involves the audience. As you see, having an accent can be a negative factor in the evaluations but not in all cases.

The Department administrators concluded that the only factor that seemed to consistently bias the TES scores was gender of the instructor. Traditionally, women professors in the Department received lower scores than men. The chair said:

Race and ethnicity and having an accent can be a factor influencing the evaluations but not in all cases. With gender it is a different situation, when we break out the TES evaluations by gender you can find significant differences between male and female professors. Women receive lower

ratings than men. The evaluations are harder for women, especially for young woman.

A member of the Executive Committee added:

People from the Office of Instructional Support and Improvement say there is literature showing that gender is not a factor in the evaluation. But, we have found that even if gender doesn't show in the global results, gender bias is there.

A member of the Executive Committee, added that although the campus had tried hard to promote diversity, women and minority faculty may be at risk.:

The university climate may not be supportive for minority faculty. We had a minority professor who left because he was looking for a university climate more supportive of diversity. He was doing just fine but he said that he wanted to look for a more diverse environment.

To make students more sensitive to issues of diversity, the Department provided them with a broad array of courses exposing students to different cultures. The Chair said:

The Department teaches a broad spectrum of students and cross-lists the courses with other Departments, such as Latino studies, Native American studies, African American, etc. We also offer teaching workshops for faculty and for our teaching assistants. One of these workshops is on classroom atmosphere with regard to difference (gender, race, etc.). We try to make faculty and teaching assistants aware of diversity but, most of all, we try to make them aware of the implications of diversity for teaching.

Our Department also offers Ethnic History courses. We have African American, Asian American and Latino courses. Usually the person who teaches these courses is from the ethnic group. What you find is that graduate students usually become interested in taking courses that are related to their race and ethnicity. So more minority students register for Ethnic History courses than do non-minority students. One of the unintended results is a concentration of students of the same culture/race in those courses, This, unfortunately, reduces the opportunities for them to interact with other students that are not of their ethnic or race.

What we don't do, but we probably should, is to look if the instructor is addressing different student backgrounds when she or he teaches.

Course sequence as a missing part of the evaluation of teaching. The Department had paid little attention to whether instructors had provided students with the required knowledge for advanced courses. Members of the Executive Committee felt that this perhaps should be an important part of the evaluation. One member of the Committee stated:

Sequence is not part of the evaluation. We don't formally look at how an instructor provides needed background for other courses. This is because we don't formally require coherence or sequence between courses. Personally, I would like to have all majors writing research papers. Many [instructors] do that but not all. There are prerequisites for our courses, like writing across the curriculum. All our tests are essay; students have to write; there are no multiple choice exams.

Another member of the Executive Committee agreed that sequential course articulation was missing from their evaluation of teaching:

I don't think sequence, the way one course prepares students for other courses, is being evaluated. It would be very difficult to do that. It could require more sophistication to evaluate this. Maybe you can look at whose students are better to handle a higher level ... Until now, this is hard to do except at an anecdotal level. You have cases of someone who does a very good job in preparing students in some areas. For example, you can have someone whom you know is lousy as a teacher but who did a good job teaching gender issues. When his students come up with interesting research questions, you realize they have developed sensitivities and a level of sophistication that is not expected of them at this level. Then you have evidence that their instructor did a good job.

Although the Executive Committee considered course to course articulation to be important, evaluating course sequence was problematic. The Department was not always able to offer courses in an ideal sequence. Professors were not always available at the right time.

The Chair said:

We should include course sequence when we evaluate teaching but for doing that we have to [offer] the courses in the sequence. Even if we think this is

important and that there is an ideal sequence of courses, sometimes we can't offer courses during certain semester because we don't always have available instructors who could teach the course that semester. It is hard to have a coherent curriculum when one third of the faculty are out doing research.

Another member of the Executive Committee added:

For most undergraduate students, course sequence and specialization is not too important. Most History majors don't become historians; they usually go into Law school or study for an MBA. There is a certain percentage of students who become History teachers. I think it's about 10%. There are others who go to graduate school to become historians but the percentage is very small.

### Summary

The College of Arts and Sciences at Midwestern University was a college with a strong reputation for its interest in improving teaching. This was reflected in the creation of teaching awards, and policies that stressed the importance of a comprehensive evaluation of teaching, and especially the creation of a Teaching Academy.

Although the College had delegated to departments the responsibility for making decisions about teaching and professional development, it stressed the importance of linking teaching and student learning and for using multiple sources of evidence when evaluating teaching. College policies also put the importance of peer evaluation over other evidence.

To provide professional development for faculty and mentoring, the College had created a Teaching Academy. This Academy had staff from the Office of Instructional Support and Improvement, which resulted in both benefits and limitations. One of the benefits was having highly trained staff knowledgeable of the evaluation process. This staff had experience providing professional development across campus. On the other hand, the

approach used by OISI followed teaching principles not fully tuned to the specific needs of individual departments.

Other professional development and evaluation activities were provided at the departmental level. This was the case of Christy's Department, the History Department of the College of Arts and Sciences. It had a reputation for good teaching. This was evident by the number of teaching awards obtained by its faculty, discussions about teaching that made new professors sensitive to the importance of teaching, and high expectations for teaching as expressed by Department administrators.

The Department was also known for scholarship and progressive political views. In some polls, it was ranked among the best History programs in the US. Its liberal orientation was known across the campus. The progressive mien was evident in department policy opposing the use of Native American symbols as university mascots, and for a teaching approach that respected the accomplishment of a diversity of peoples throughout History.

The Department's cultural orientation for teaching History resulted in some trade-offs. In coming to this orientation, the Department to some extent gave up generalization and comparison for complexity and understanding of diverse perspectives. This contrasted with the State's secondary school emphasis on right answers to complex Historical questions.

The members of the Executive Committee and the Chair spoke of good teaching as supportive of a scholarly and cultural orientation, and stressed the importance for instructors to adjust their teaching to contextual needs. They saw good teaching as promoting citizenship, understanding of diversity, respect for students as persons, sparking their curiosity and as helping students to make connections with the content. Christy's teaching

philosophy and style fit well within this approach, although it is hard to point to how that fit entered the formula.

More straightforward, teaching was evaluated using two main sources of evidence, student ratings and peer reviews. Student ratings were obtained from the campus evaluation system (TES) and from a Departmental survey administered to students enrolled in History courses at the end of the semester. Peer evaluations of teaching took place every semester and were conducted by selected members of the Executive Committee. The professors making these observations wrote a report for both the Administration and the instructor. It was the expectation that these reports were going to be used for both instructional improvement and as evidence of good teaching. Other evidence when evaluating teaching included student projects, course materials and examinations, the instructor's teaching philosophy, and so on.

The Department had recognized both strengths and limitations in using the campus TES ratings of instruction and the peer evaluation system. Student ratings were found to be useful in providing evidence of teaching for general courses with a medium class size of 30- 40 students. Student comments made on the forms were said to be especially valuable for some professors in identifying what went wrong in their teaching and what could be improved. On the other hand, TES were seen as sometimes influenced by extraneous factors, such as the instructor's interest in instructional technology, the rapport between the instructor and students, student grade expectations, class size, and by the race, gender and age of the instructor. Moreover, TES ratings were said not to be good indicators teaching quality since instructors who covered little content were able to obtain high scores. Some instructors were dismayed that the ratings did not reflect what students learned in the course.

Peer evaluations were considered very useful and were encouraged by the Department. They were said to be especially valuable for instructional improvement and for interpreting TES results. They brought in the contexts of teaching and were used to identify instructor strengths for teaching different kinds of courses. The main problem with this evaluation source was the conflict of interest of the observers, since the peers played a dual role as evaluators, reviewers for salary increase and promotion, and colleagues. Some professors said that some observers had problems in identifying instructional quality. Another problem associated with peer observations was that they seemed to be influenced by TES ratings and by the instructor's reputation.

The documentary evidence about teaching was placed in the professor's teaching portfolio. Members of the Executive Committee and the Chair reviewed this portfolio when making administrative decisions, such as annual and third year reviews for tenure and promotion. Procedurally, in these reviews, the Committee made a review of the instructor's scholarly production before looking at teaching. If the instructor scholarly production was high, the Committee proceeded to review the evidence of teaching. In other cases, the review of teaching was more comprehensive. They reviewed first the information from the TES ratings and used the peer reports to bring contextual information for understanding the scores. If TES scores were lower than 2.5 (on a scale of 5), the professor was asked to develop an improvement plan. It could involve mentoring, observing other faculty teaching, or searching for help in one of the campus units that provided consulting on teaching improvement. Helping another faculty member to improve was not rewarded as teaching, it was considered a service to the Department.

Those who obtained good ratings were rewarded with departmental, campus and college awards. Those who did not improve were affected mostly in their reputation as teachers. This could affect their chances for getting tenure if their research production were low. In general, research had priority over teaching. Still, most committee members indicated the role of teaching had increased in administrative decisions. Teaching was said to be especially important in cases of poor scholarship, then it became the one strong reason for granting or denying tenure.

Some of the pressures identified at campus and department levels were present in this History Department. One pressure was that for measuring student outcomes. An ad-hoc committee struggled to find meaningful ways of assessing student outcomes. Some of the professors were concerned about the problems of assessing teaching through student outcomes.

Another emerging issue from the evaluation of teaching in this department was the possible bias of students when evaluating women faculty and instructors of diverse race and ethnicity. The members of the Executive Committee here found that, over the last 10 years, student evaluations of teaching were lower for women, especially young woman. They believed that race and ethnicity also, as well as any kind of physical or cultural difference, such as instructor accent, might negatively influence student evaluations of teaching. They had conducted inquiries. When a claim was made in this sense, they set out to learn if it were a case of bias or a case of ineffective teaching.

In general, there was not a conflict between the Campus and Department evaluation. The Department evaluation system supplemented the TES evaluations and provided

information about the context of teaching. The Department was sensitive to issues of student diversity but did not include them as part of the evaluation of teaching. Sequential articulation from course to course was identified as very important and some members of the committee urged good articulation be included in the evaluation of teaching,

Evaluation results were used for both instructional improvement and for making administrative decisions. The Department claimed the main use was for helping make the instructor's case for tenure. Department support for improving teaching was primarily through its mentoring system. Although, for instructors like Christie who did not obtain TES scores below average and scored above average in her required undergraduate courses, there was less support provided for improving teaching. It is so described in the following chapter.