1992

What Tenure Files Can Reveal to Us About Evaluation of Teaching Practices: Implications for Instructional/Faculty Developers

David Way

What Tenure Files Can Reveal to Us About Evaluation of Teaching Practices: Implications for Instructional/Faculty Developers

David Way

Cornell University

This article describes a study that examined the way teaching is evaluated for tenure and promotion at Cornell University. After reviewing tenure files from several different colleges in the university and interviewing deans in the colleges, the author reported great variation among colleges in the kind of information that is collected and retained in tenure files. In addition to providing the results of the analysis of the files, the author identifies recommendations that were made based on the findings and discusses the role of instructional/faculty developers in assisting faculty in the documentation of their teaching.

Over the years POD members have suggested that instructional/faculty developers stay clear of the summative evaluation process. Summative evaluation of teaching is carried out in order to make personnel (tenure and promotion) decisions, whereas formative evaluation is undertaken for the improvement of instructional practice. This distinction is extremely important because, it has been argued, a confounding of these two purposes by instructional/faculty developers creates a conflict of interest. As the director of a center for instructional support, I see my services benefiting the faculty

and ancillary teaching staff (instructors, lecturers, teaching assistants). To that end, my staff and I engage in formative evaluation.

This article, describing a study carried out at Cornell during the last year shows, however, that the formative and summative evaluation functions can be linked programmatically—to the advantage of everyone—including administrators, faculty and instructional development staff alike. During the past year I have collaborated with the Dean of Faculty and several tenured faculty members selected by the dean to examine the way teaching is evaluated for tenure and promotion at Cornell. Specifically, our charge was to examine "the practices and procedures used by the colleges and schools in evaluating a faculty member's teaching contributions as part of the tenure and promotion process and to recommend minimal standards for such reviews. Accordingly, the committee:

1. reviewed the literature on the evaluating of teaching to gain an overview of what has been learned from research and experimentation elsewhere;
2. learned how each of Cornell's colleges and schools evaluates a faculty member's teaching of undergraduate and graduate students; examined tenure files from the schools and colleges; and interviewed each dean;
3. recommended procedures and guidelines to help ensure proper assessment of a tenure or promotion candidate's qualifications as a teacher."[1]

What I would like to do in this article is to describe the method I used as a major participant in this project, explain the results of an analysis of 19 tenure files selected from across disciplines, identify the outcome of the project, and discuss the implications of that product for instructional/faculty developers.

Method

Tenure File Review[2]

The Dean of the Faculty asked the deans' permission to look at tenure files in each of the colleges. Nineteen tenure files were selected from seven colleges and schools for analysis. A selection was made from a pool of faculty who were granted tenure in the period from July 1, 1987, through June 30, 1991. The tenure files reviewed were chosen according to the following

---

[1] Based on a report of the Select Committee, *Evaluation and Recognition of Teaching.* Cornell University, Ithaca, N.Y. Jan. 14, 1992, p. 2.

[2] *Ibid,* Appendices, p. 17.

criteria: tenure had been granted within the previous three years, the candidate had at least a 45% responsibility to teach as indicated by the chair, and the broadest number of colleges and departments were represented in the sample.

The review of tenure file contents was based on the assumption that a tenure file should provide thorough documentation of the process by which the candidate was evaluated, both in terms of research and teaching. It is recognized that there is much more to the tenuring process than what can be put in a file. However, the file is the primary document used to make the candidate's argument for tenure and should therefore reflect a degree of thoroughness and detail sufficient to stand on its own as a source of evidence.

Anonymity of both departments and colleges was maintained in this study to avoid an accusatory tone in the report and to support an atmosphere of open, constructive debate. Discussion of file contents was kept very general. Specific supporting evidence was reported with minimal identification of department or college. Because the discussion of file contents is based on such a small sample of files, generalizations are limited. Conclusions in this section are meant to provide an overall picture of central tendency while acknowledging that there are many exceptions to that tendency.

After a review of 19 tenure files, 20 items were identified as the major teaching dossier components (Seldin, 1991). These 20 items were then organized into the following 5 categories:

Process Data:
- appointment letter
- curriculum vita
- reappointment letters/annual performance reviews
- job description
- personnel forms (% teaching vs. % research, service, extension)
- letters (from chair, dean) soliciting evaluation of the candidate by students, peers

Peer Review Data:
- *ad hoc* committee report
- chair report
- reviewers' letters
- observational reports
- letter from the dean

Student Evaluation Data:
- student evaluations of instruction
- undergraduate letters of recommendation
- graduate student letters of recommendation
- alumni letters of recommendation
- advisee letters of recommendation

Teaching Materials: • course materials
                        • teaching itinerary (name and number of courses
                           taught by semester/year)
Self Evaluation Data: • personal statement
                        • teaching improvement strategies/activities

The five categories were selected because they define the major data sources available to faculty and administrators who do the evaluating—students, candidate, peers, administrators—and because they provide a framework with which to match the appropriate evaluation source with the aspect of teaching that is being evaluated. For example, students are an appropriate evaluation source for classroom teaching quality, whereas peers are an appropriate evaluation source for course design and content, while the candidate is a rich evaluation source for the effectiveness of teaching improvement strategies. These five categories will serve to frame the discussion of the tenure file contents in the results section below.

Interviews with the Deans[3]

In addition to information obtained from the analysis of the files, other contextual information regarding the tenure process was obtained through interviews conducted with the deans from June 9 through July 10, 1991. The questions that formed the basis of the deans' interviews are listed in the appendix. In general, the deans were very cooperative and eager to discuss the tenure system.

Results

Tenure File Analysis

In general, inclusion of teaching dossier components varied considerably by college. The number of colleges that had each dossier-component in tenure files is shown in Figure 1.

Process data. In all cases documentation of the tenure process was very thorough. However, there was considerable variation in the way the communication between the candidate, chair, reviewers, *ad hoc* committee, central administration, and students was documented and organized. An important factor in evaluating the communication of the tenure process (Seldin, 1984)

---

[3] Ibid, Appendices pg. 25.

FIGURE 1: Number of Colleges With Each Dossier Component

Number of Colleges
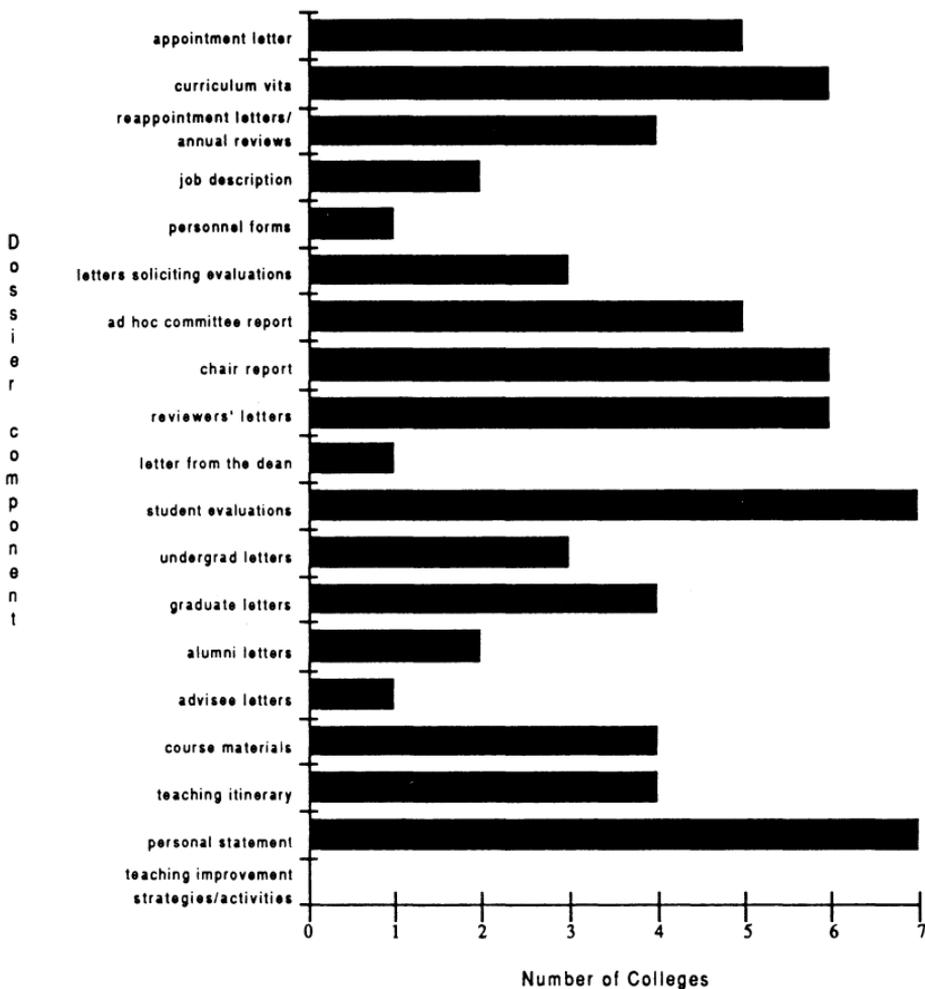
is the explicitness of the task the reviewers (peers and students) are asked to perform. Letters from chairs (both departmental and *ad hoc* committee chairs) soliciting reviewers' comments ranged from very general requests to very specific lists of instructions that clarified the criteria by which the reviewer was to evaluate the candidate's teaching (either from direct observation and experience or through review of teaching materials.) An example of an explicit request from a chair that helped focus the reviewer is, "Does the subject matter presented in the individual's courses keep pace with developments in his or her field? What is the evidence?" This kind of explicitness was very rare, however, and was evident in only one of the 19 files.

The lack of specific criteria for guiding reviewers' evaluations of teaching was evident when teaching evaluations were compared to analogous reviews of research and publications. The general sense was that chairs and reviewers were much more attuned to discussing research and publications than teaching. Where colleagues were familiar and comfortable intellectually with the context of the candidate's research and published work, lengthy and thorough evaluations and discussions were the norm. In general the equivalent discussion of the candidate's teaching quality was much shallower, as evidenced in the following quotes from peer reviews of teaching quality:

Internal Reviewer: "I have heard [candidate] speak on several occasions, and he has always been lucid. His enthusiasm also comes through in his talks and it is contagious. I expect that he is an excellent lecturer."

External Reviewer: "I have heard him explain his results and discuss them in seminars. He is very clear and does not spend time on fluff. I have not had the experience of attending his classes."

External Reviewer: "He conveys enthusiasm and is well prepared in conference and seminar presentations, so I would expect he is a good teacher."

The quotes were in a file that had no course materials or letters evaluating teaching quality by any students (undergraduate or graduate.) This example is typical of a case where the primary evidence for evaluating teaching quality was student ratings scores, followed by the very general and shallow arguments of peers as evidenced in the quotes above. In spite of the fact that the majority of colleges typically included copies of appointment and reappointment letters, these documents were missing in a surprising number of files.

Peer review data. Peer review of teaching quality was much less rigorous than that of research. The level of the discussion was often set by the department chair in annual performance review letters to the candidate:

We discussed your teaching activities and my indications are that your

teaching . . . is going well and that it is well received by your students. We hope that the renovations of the laboratories . . . will provide opportunities for further improvements in our . . . teaching program.

Peer evaluation of classroom performance was very inconsistent among colleges and did not exhibit the kind of rigor suggested by the literature on the subject (French-Lazovik, 1981; Miller, 1987; Seldin, 1990; Sell & Chism, 1988; Whitman & Weiss, 1982). The minimal degree to which internal reviewers were prepared to observe and comment meaningfully on each other's teaching is evident in the following quotes that reflect a common theme among the files reviewed.

Peer letter, 9/28/89:

I regret that conflicts have prevented my attendance, since by all reports it [teaching] is excellent. My personal observation of his teaching has been restricted to slipping in at the end of his lectures . . . what I have heard has been clear, accurate and well organized.

Department head letter, 10/13/89:

I have no direct knowledge of Dr. X's performance and effectiveness as a teacher; however, judging from comments of my graduate students, his [course title] is excellent and has become a favorite . . .

Student evaluation data. Without question, student quantitative evaluation of instruction through end-of-the-semester summative questionnaires was the most prevalent data source from students on candidates' teaching quality. Letters from students—undergraduate, graduate, advisees and alumni—were the next most frequent source of student evaluation data. The advantage of numerical questionnaire data is that they preserve the anonymity of the student, which is a significant factor while students are still enrolled in school and thus in the "power web" of the candidate, as one dean described it. The advantage of the written letters is the quality of the data they can provide: writing a letter provides a student with a more open format to evaluate the candidate.

In one case the department chair took the time to synthesize the relevant comments from undergraduate student reviewers who were asked to write letters of recommendation. This process was very useful considering there were 88 undergraduate letters of recommendation in the file, but it was the only file with such thorough care given to a discussion of the candidate's teaching. It stands as an exemplary case of how a chair summarized student evaluation letters as evident in the following quotes:

Undergraduates uniformly describe him as an unusually effective, conscientious, enthusiastic teacher who enables students to do their best work, master difficult subject-matter, and gain confidence in their own intellectual abilities.

This [student quote from a review letter] clear and convincing testimony describes the experience of all the students who wrote to us from the courses he taught in spring 1988 and in fall 1989. Since the most disturbing aspect of some of the student responses two years ago was the suggestion that he could be authoritarian and coercive in his teaching, we are reassured by all these letters which suggest precisely the opposite.

It seems clear that like many young assistant professors [candidate] was too demanding in his first dealings with graduate students, imposing admirable but often excessive standards of professionalism both in the classroom and as a special committee member, and expecting his students to share his commitment to his own projects. As the letter from [student] suggests, however, he has since become more realistic and flexible. And all the letters attest that he is always extremely conscientious and helpful.

One should conclude, I think that [candidate] is an intellectually stimulating and enabling graduate teacher, with an expertise and commitment that many of our students find particularly valuable, one who has had trouble finding the appropriate mode in which to exercise authority, but who has now learned to do so.

Where quantitative data were provided, there was great variation in the form in which they were presented. The kind of guiding principles for collecting and reporting such data that are suggested in the literature (i.e., do not carry scores beyond one decimal place, sample over time and course format) were little in evidence in the files reviewed.

Surprisingly, one major source of information from students, the products of their work with the candidate such as exams, papers (both good and bad) and grade sets from year to year, were rarely included in the files. This lack stood in marked contrast to the extensive volumes of reproduced published articles that made up the bulk of many of the files.

Teaching materials. The literature on the evaluation of teaching for tenure and promotion mentions the inclusion by the candidate of course teaching materials (syllabi, exams, reading lists, teaching plans and notes, computer software, etc.) as a way for peer reviewers, both internal and external, to evaluate instructional design skills even though they may never have seen the candidate teach in the classroom.

The inclusion of teaching materials in the files that were reviewed was

inconsistent among departments and colleges. Some files limited this data source to a "teaching itinerary" that consisted of nothing more than the candidate's teaching schedule: what courses were taught each semester, from year to year. In the files reviewed, the most prevalent course materials were syllabi, reading lists, and ungraded exams. However, only three out of 19 files contained teaching materials of any kind. This number seems out of balance considering the extent of publications in the tenure file.

Although a "teaching itinerary" may be useful to someone in the department who has taught the same courses, it is not sufficient to evaluate a candidate's instructional design skills, which include preparing learning objectives, creating effective instructional experiences—strategies to sequence, interpret, and present information—and designing tests and graded assignments. An activity as challenging and complex as teaching deserves more thorough documentation than a list of courses taught every semester. Examples of teaching materials suggested by the literature for inclusion in a teaching dossier include syllabi, reading lists, handouts, non-print materials (or a description of them), problem sets, assignments, graded exams (both good and bad), research papers and final projects, final grade distributions, and examples of written feedback to students on their work.

Self evaluation data. Information written and prepared by the candidate providing self evaluation data is mentioned repeatedly in the literature and was prevalent in many of the files reviewed. Some kind of "personal statement" dealing with research and teaching was included in 13 out of the 19 files. Once again, the depth of thinking reflected in the personal statements reviewed about research was greater than that devoted to teaching. This disparity was clearly evident in the amount of space devoted to describing research goals and agendas compared to that devoted to teaching goals and agendas.

Various authors have provided suggestions and guidelines for candidates for tenure to describe their efforts and activities to improve teaching (French-Lazovik, 1981; Miller, 1987; Seldin & Associates, 1990; Sell & Chism, 1988; Whitman & Weiss, 1982). However, the kind of discussion these authors have suggested was not evident in the files reviewed. In one case the college had a multi-page form developed for the candidate to fill out for his or her annual performance review that requested specific information about teaching goals, improvement strategies and activities. The candidate in question, however, had done a cursory job of filling it out. This superficial approach was noted by the candidate's supervisor to the associate dean in the annual review documentation, but it did not prevent the candidate from being granted tenure.

Deans' Interviews

Quality of file information. The deans had mixed responses to the question of overall file quality and inclusiveness. The degree to which they were satisfied was strongly influenced by the degree to which guidelines for file contents were clearly articulated and adhered to by department chairs. Whether the guidelines were adhered to properly was a judgment made at the dean's level; and when there was insufficient information, in whatever form, deans exercised their authority to return the file for supplementation. However, whenever this action was necessary (a situation that was mentioned by the majority of deans), it had the undesirable effect of slowing down the process.

One dean said that the quality of teaching documentation was not as in-depth, helpful or well-conceived as that devoted to research. Another emphasized that *quality* of information was a more preferable guide for inclusion in a file than *quantity* of information, especially in terms of published research. Two areas of information mentioned to be in general short supply were course materials and alumni reviews. The deans were unanimous in their reliance on departmental chairs for producing good promotional packages and tenure dossiers.

Teaching dossier contents. In discussing the items that should be included in tenure and promotion dossiers, the deans echoed some of the shortcomings identified in the file analysis. In terms of personal statements, one dean felt strongly that a detailed essay (with a length limit) on course design and instructional delivery with an emphasis on development of practice should be the norm. Another felt dossiers should contain more course materials, whereas another felt better classroom observation data should be included. Several deans emphasized the importance of having someone (department chair or *ad hoc* committee chair) provide a summary of detailed student comments of at least half the candidate's courses, the courses to be included to be chosen by the *ad hoc* committee.

An important issue in terms of dossier contents has to do with who is responsible for building the dossier—the candidate or the department chair. Better dossiers will require clear guidelines as to data sources, evaluation criteria for all data sources, and a clear explanation of who is responsible for assembling the dossier. One dean felt that faculty should be made aware as part of their appointment that they are responsible for providing documentation of teaching.

Another dean raised the issue of anonymity of student evaluation data. The fact that letters of recommendation in their original form (the prevalent form from the file analysis) do not preserve the anonymity of the student is

an important consideration in judging the credibility of the student's evaluation. While students, either undergraduate or graduate, are still working with the candidate, they are in what one dean called the candidate's "power web." This situation may prevent students from being as candid in their written remarks if they know they may be identified at some point by someone in the tenure decision process. One dean suggested having someone (department or *ad hoc* committee chair) summarize student letters.

Overall process. In general, the deans felt that the tenure and promotion process needs to become more efficient in terms of time and other costs, yet more rigorous. Some authors (Sell & Chism, 1988) have advocated the utility of evaluation forms for evaluating both tenure files and classroom instruction. One dean said that the hiring and appointment phase of the process is a more important milestone influencing the quality of teaching than the tenure milestone. Another dean felt that, in general, the deans have to be willing to take more risk rewarding excellence in teaching because they have more power to do so.

Study Outcome

As a result of the study a report was written that included several recommendations. It has been distributed to the faculty to solicit their reactions through an open exchange. The committee recommended that:

- it should be a matter of policy that colleges and departments treat and reward teaching and research on an equivalent basis;
- departments and colleges, consistent with university policy, express equal value for teaching and research by maintaining and supporting the highest standards in their evaluation of both these activities;
- rewards should be developed and policies designed for excellence in teaching and for significant instructional development, including salary-based as well as non-salary based incentives (e.g., additional TA support, teaching aids, etc.) in direct support of teaching activities and should be designed to affect the largest number of faculty members;
- all departments (or other appropriate units) establish a standing committee on teaching whose members would be responsible for overseeing peer evaluation of a tenure candidate's teaching;
- each college should have and enforce specific guidelines governing what must be submitted about teaching in the file of a tenure candidate;
- each college dean establish, maintain, and monitor guidelines for obtaining student course evaluation materials (i.e., questionnaires, solicited letters from students) and the way these materials are reported;
- a teaching evaluation handbook be developed and made available to all

colleges and departments in order to encourage consistency in the evaluation of teaching;

• the provost take appropriate steps to set in motion the recommendations embodied in this report.[4]

Preliminary reactions by central administrators and several faculty members have been encouraging in terms of support for the recommendations. The objective is to submit the report to the provost, after the faculty's reactions have been heard and assimilated, for his compliance with the final recommendation above.

Discussion—The Role of Instructional Developers in the Tenure Process

There is no one way a faculty member should take advantage of instructional development services; it is best handled on a case-by-case basis. Within this latitude, however, instructional developers can review the contents of a teaching dossier with the faculty member prior to submission for tenure review to ensure completeness, quality of data, and appropriateness of format. Not only can this "quality control" practice potentially increase a faculty member's confidence in dossier quality, but it can aid department heads and deans by providing an independent quality control point to minimize the rejection of files because of poorly documented teaching content.

As a resource to (rather than as an advocate for) faculty or administrators within the tenure process, instructional/faculty developers are in a unique position to assist in the data gathering and documentation of instructional development for tenure files. As part of our normal operations, we can serve faculty in three major developmental areas: instructional delivery skills, instructional design skills, and activities associated with general innovative developments.

Activities in all these areas are valid for inclusion in a tenure file as a means of documenting relative teaching quality and development. It is important, however, that the primary beneficiary of instructional development services be the faculty member. Department heads and deans can benefit secondarily through improved data to support the improvement of teaching, but instructional development exists to serve faculty. If instructional development centers are perceived as primarily serving administrators

---

[4] Ibid., pgs. 7-8.

for the tenure system, there exists a conflict of interest. Instructional developers can serve both individual faculty and administrators within the tenure process by continuing to orient their services to faculty, but with the understanding that the faculty member can and should describe in a personal statement the strategies and activities engaged in to improve instructional practice. Instructional development center staff can assist faculty members in all of these steps.

Conclusion

Although departments can review, evaluate and recommend where teaching must be enhanced, instructional/faculty developers provide faculty and TAs with the resources and expertise necessary to improve practice and to document that improvement. Although department heads, deans and the provost are in a position to enforce comprehensive evaluation criteria and guidelines for tenure file development and maintenance, the burden of adequate documentation rests with the individual faculty member. Instructional/faculty developers serve the faculty member as a resource to ensure adequate documentation of the improvement in teaching practice. If proof of teaching development/improvement is a criterion for tenure and promotion decisions, the formative evaluation process has begun to inform positively the summative evaluation process.

References

French-Lazovik, G. (1981). Peer review. Documentary evidence in the evaluation of teaching. In J. Millman (Ed.), *Handbook of Teacher Evaluation.* Beverly Hills, CA: Sage Publications.

Miller, R. (1987). *Evaluating faculty for promotion and tenure.* San Francisco: Jossey-Bass.

Seldin, P. (1984). *Changing practices in faculty evaluation.* San Francisco: Jossey-Bass.

Seldin, P., & Associates. (1990). *How administrators can improve teaching.* San Francisco: Jossey-Bass.

Seldin, P. (1991). *The teaching portfolio - A practical guide to improved performance and promotion/tenure decisions:* Bolton, MA: Anker Publishing.

Sell, G. R., & Chism, N. (1988). *Assessing teaching effectiveness for promotion and tenure: A compendium of reference materials.* Columbus, OH: Center for Teaching Excellence, The Ohio State University.

Whitman, N., & Weiss, E. (1982). *Faculty evaluation: The use of explicit criteria for promotion, retention and tenure.* ASHE-ERIC/Higher Education Research Report, No. 2. Washington, DC: The George Washington University, School of Education and Human Development.

Appendix
Questions for Interviews with Deans

1. When a tenure dossier is presented for your consideration are you satisfied with the quality of the information you need to render a decision about the candidate? What additional information would you have liked to have had to make whatever action you are required to take clearer, more defensible?

    Does the quality of the information provided differ by field or department?

    How significant is the department chair's role in influencing your decision?

    Faculty are expected to vote on each tenure candidate, and in some colleges not only is the vote recorded, but faculty members are asked to provide a rationale for the vote cast.

    How does the final tally influence your decision?

2. How significant is the recommendation of the ad hoc committee in helping you to decide?

    What are the principal considerations involved in selecting the members of an ad hoc committee?

    In general do they provide you with new information or insights?

    What value do they provide to the tenure process?

3. How do you evaluate the candidate's contributions in teaching undergraduate and graduate students? [Is it common practice for the dossier to include letters solicited from current and former graduate students, course evaluations from undergraduate courses, etc.?]

    Is the information about the candidate's abilities and contributions as a teacher sufficient for your purposes?

    What additional information, if any, should be included in order for you to render a judgment in these matters?

    How important are course evaluations? Could they be made more useful to all the parties involved in a tenure decision?

    How does the information contained in the graduate student "letters" compare with undergraduate course evaluations in terms of the influence they may have on your decision?

4. Why do dossiers contain considerably more information about the candidate's research performance than about teaching?

    How would you compare the level of effort the faculty give to evaluation of the candidate's teaching expertise and performance with that given to research?

5. If teaching could be evaluated as carefully and thoroughly as research, what difference do you think it would have upon the outcome of the tenure review process?

6. What influence do you have on the nature and quality of the information that is collected and presented to the faculty and yourself?

7. If more time and effort on the part of the faculty were required to evaluate the teaching performance of a tenure candidate, do you think such efforts would be endorsed by the faculty?

8. Are your college's guidelines (for carrying out a tenure review) strictly followed by the department in the materials submitted to you?

9. Should a dean be willing to recommend a faculty member for tenure based primarily upon an outstanding record of performance in teaching? Have you ever had the occasion to do so?

10. Does the university provide appropriate rewards for excellence in teaching? What kinds of rewards would best serve this purpose?