Peer Review of Teaching
Warrington College of Business Administration
University of Florida

Prepared by
Warrington College of Business Administration
Teaching Committee

First Draft: Spring 1998
Final Draft: Fall 2001
Adopted: March 2002 in College Faculty Meeting

Each College has been charged with developing a peer teaching review plan that follows guidelines laid out in *Recommendations for Peer Review of Teaching at the University of Florida,* which was produced by a UF task force in 1995. If we disregard our obligation and do not formulate our own peer review plan, the Teaching Committee fears that the University will impose a less desirable peer review plan upon us. In either case, the University will require some peer review process.

Peer review provides a valuable opportunity to supplement the information gleaned from student teacher evaluations about a faculty member's teaching effectiveness. For example, peer review may reveal that a faculty member is teaching a very demanding and ultimately useful course that is not fully appreciated by students now in the course. Admittedly, peer review of teaching can involve a lot of 'noise', as the empirical studies cited below suggest. But the Committee has concluded that similar issues arise in evaluating research. We are convinced that there is much to be gained from making the most of this opportunity to judge teaching not solely by student teacher evaluations.

Our recommendations for implementing a peer review plan in the College are first summarized and then described in detail. The UF report and literature that form the basis for these recommendations are surveyed at the end of our report.

Executive Summary of Teaching Committee Recommendations

An evaluation of a faculty member's teaching effectiveness is required for the third year review, for the tenure decision in typically the sixth year, and every seven years for the post-tenure review, as well as for promotion to Professor and for teaching awards. Peer reviews will take place at these times, but can be done more frequently if requested by the faculty member or by a unit head who is concerned about the faculty member's teaching. Peer reviews will be done by a committee of three faculty, who will rely on an examination of the faculty member's teaching portfolio and on classroom observation. Peer reviews also should offer suggestions for improving teaching. We also recommend that the University Center for Excellence in Teaching

be given enough funding so that it can be staffed with trained teaching consultants, who can provide help on becoming a more effective teacher.

Recommendations for Implementing a Peer Review Plan in the College

The UF report focuses on the *summative evaluation of teaching performance*, which seeks to determine whether the faculty member is doing a good job in the classroom. This information is needed when the faculty member comes up for third year review, tenure, promotion, or post-tenure review or when he or she is a candidate for a teaching award. Information about teaching quality also plays a crucial role in determining raises. Consistency with the UF Task Force Report requires that any College peer review plan include a summative evaluation of faculty teaching quality that addresses significance of content and level of pedagogical expertise, considers multiple measures that include classroom observation, and uses multiple evaluators.

It is hard to critique someone's teaching without coming up with recommendations for improvement.   We recommend that the judgment of teaching effectiveness (i.e., the summative evaluation) be supplemented as much as possible with a *formative evaluation of teaching*, which provides feedback to the individual faculty member for the purpose of improving teaching performance.

We recognize that formative evaluation often would be more valuable if observation of a faculty member's teaching by fellow faculty were supplemented with observation by trained teaching consultants. For example, in the Stern Business School at NYU, teaching consultants offer advice on improving teaching that is based on classroom observation (sometimes recorded on videotape) and/or conversations with students. This resource is needed to assist faculty who are struggling with their teaching and would be a great help to faculty who wish to become as effective as possible. We strongly recommend that the University Center for Excellence in Teaching (UCET) be given enough resources so that faculty can easily take advantage of this help. UCET seems to operate on a shoestring budget. It does not provide equipment or a videographer to tape a lecture. Instead, faculty are expected to obtain a camcorder and tripod from the Office of Instructional Resources and to find someone to tape their lecture. Faculty then are asked to bring the videotape of their teaching to UCET for evaluation. UCET's Director, Constance Shehan, reviews these if she has the time and passes this task on to faculty members on UCET's Board otherwise. We believe that UCET needs the equipment, videographer, and dedicated teaching consultants to make getting this help as easy and valuable as possible. New faculty should be made aware of all the campus resources for assisting them to become better teachers.

Proposals regarding the method, content, and timing of peer reviews are outlined below.

Proposed Method

The peer review committee will consist of three faculty members, two members from within the same department and one member from outside the department. The unit head, after consultation with the reviewee, will appoint the review committee; at least one member of the committee must be selected by the reviewee. The peer review committee may have the same members and

operate concurrently with another review committee (e.g., three-year review or post-tenure review).

The peer review committee will evaluate a faculty member's teaching quality in terms of significance of course content and pedagogical effectiveness through examination of the teaching portfolio and classroom observation (in-person or viewing of on-line materials for internet and TV-Replay classes). The reviewee's annual teaching portfolio should be considered sufficient for the committee's non-classroom evaluation. If the teaching portfolio is deemed to provide inadequate evidence on the reviewee's teaching quality, the unit head can require the reviewee to revise the portfolio. The reviewee should be allowed to select the timing of the in-person classroom observation, subject to reasonable constraints. At least two of the three committee members must participate in the classroom observation.

The peer review committee will provide the reviewee with a copy of the peer review report and will provide the reviewee with the opportunity to meet with the committee and to respond to the draft report. After any meeting with the reviewee and/or receiving the reviewee's response, the committee will provide the unit head with a single report on the assessed quality of the reviewee. The reviewee may attach a response if he or she wishes. This report and any response will be included in the reviewee's personnel file.

Report Content

There is no required format of the peer review report (or a particular checklist for evaluation of the reviewee's teaching portfolio or classroom observation). The report should address the following aspects of the reviewee's teaching performance.

1. mastery of course content
2. selection of course content
3. course organization
4. appropriateness of course objectives
5. appropriateness of course materials (such as readings, media)
6. appropriateness of evaluative devices (such as exams, written assignments reports)
7. appropriateness of methodology used to teach specific content areas
8. commitment to teaching and concern for student learning
9. student achievement

In addition, the report should provide specific suggestions for improving teaching where particular weaknesses are identified and indicate what resources are available for the reviewee.

Timing

A *summative* teaching evaluation is required for the third year review, for the tenure decision in typically the sixth year, and every seven years for the post-tenure review, as well as for promotion to Professor and for teaching awards. It is expected that peer reviews will take place at these times and will be coordinated with other aspects of the faculty review process. In addition, a unit head may require more frequent peer reviews of any faculty member for whom

there is a concern about teaching, and a faculty member may request a *summative* peer review. Given this schedule, we see no need to require more frequent peer reviews.

Similarly, a faculty member may request a *formative* evaluation of his or her teaching, either by peers or by trained teaching consultants in UCET. The faculty member may be more candid in describing teaching weaknesses and in accepting constructive criticism in this setting, since promotions and raises are not on the line that year. This is especially recommended for tenured faculty sometime in the seven-year interval between summative peer evaluations.

Background Material

Summary of UF Task Force Report

The UF task force produced nine recommendations to assist departments and colleges in developing their own peer review plans.

1. At a minimum, the peer review process should include an evaluation of the teaching portfolio and classroom visitation.
2. The faculty member should participate in the selection of the peer review panel. The peer review process should define 'peer' and the peer review panel should include at least three faculty members, with at least one member from outside the department.
3. The peer review process should include both summative and formative components.
4. The peer review process should specify the frequency of summative peer review. The process is mandatory for tenure and promotion decisions and teaching award candidates. At a minimum, untenured faculty should be reviewed two to three years before the tenure evaluation and the year of the tenure decision. All other faculty should be reviewed no less than once every five years. Additionally, faculty should be able to request a peer review.
5. Peer review plans should specify the content of peer review reports. The reports should be considered similar to reports of the peer review of research. The reports should use multiple variables in the evaluation of teaching: course design, classroom observation, teacher/student interaction, student evaluations, student performance, and instructional constraints.
6. The peer review plan should specify the structure and process of classroom observation. Multiple observers should evaluate classroom performance on different occasions.
7. The peer review plan should specify the process for preparation of the report. Three alternatives are suggested: (a) three panelists prepare individual review letters, with all placed in file, (b) three panelists prepare letters independently then meet to construct a single letter for the file, or (c) a combination of (a) and (b).
8. The faculty member should receive a copy of the letter placed in the file, and if requested, a meeting between the peer review panel and the faculty member should be held.
9. Staff development should be provided by colleges and/or departments (with the assistance of the University Center for Excellence in Teaching) for faculty serving on peer review panels (e.g., training, a written manual, sample review letters, written manual, or instructional videotape).

Peer Review of Teaching Issues

In reviewing the literature on the peer review of teaching, several issues consistently appear. These are addressed briefly below. Some of these issues relate primarily to *student* evaluation of teaching, but are included here because issues surrounding peer review of teaching are necessarily interrelated with student evaluations.

Students as Evaluators

Studies have investigated the correlation between teacher evaluations and amount learned, as measured by a common exam, when different instructors teach multiple sections of the same course. Adjusting for potential differences in ability across sections, meta-analysis studies have shown that the correlation between student ratings and exam measured achievement average about .40 (Abrami, Cohen, and d'Apollonia 1988; d'Apollonia and Abrami 1997).

Faculty/Administrators as Evaluators

Colleagues' and administrators' ratings of instructors are not correlated with student evaluation of instructors nor with other indicators of teaching effectiveness (e.g., achievement) (Centra 1979; Koon and Murray 1996; Marsh 1987; Murray 1980). Moreover, ratings provided by colleagues and administrators do not correlate with each other (Howard, Conway, and Maxwell 1985).

Trained External Observers as Evaluators

Trained external observers can accurately differentiate between teachers that promote high, medium and low achievement. Murray (1983) found that ratings based on 18-24 observations (videotaped lectures) could predict teacher effectiveness, but only when the ratings of multiple trained observers were averaged. Ratings by individual observers did not correlate with indicators of teaching effectiveness.

What Should Be Evaluated?

Marsh and Roche (1997) review the literature on student evaluations of teaching effectiveness and find support for nine factors. Cohen (1987) used 41 studies to calculate correlations between each factor and achievement (measured by test scores):

Dimension of Teaching Effectiveness	r with Achievement
1. Learning/Value	.39
2. Instructor Enthusiasm (Stimulation)	.15
3. Organization/Clarity	.55
4. Group Interaction	.52
5. Individual Rapport (Available)	.32
6. Breadth of Coverage (Knowledge)	.50
7. Examinations/Grading (Fair Evaluation)	.30
8. Assignments/Readings	.30

    9.   Workload/Difficulty   -.04

These factors have been supported by over 30 published empirical studies and are discussed in detail by March and Roche (1997) and d' Apollonia and Abrami (1997).

Biases

Is there a class size bias? It is possible that larger class sizes create a more impersonal environment and hurt ratings of teaching effectiveness. Class size has been shown to be correlated with ratings of Group Interaction and Individual Rapport, but uncorrelated with the other seven indicators of teaching effectiveness. The amount of influence class size has on an overall rating of teacher effectiveness is a function of the weight students place on Group Interaction and Individual Rapport relative to the other seven factors.

Is there an expected grade bias? It is possible that evaluations of instructors can increase or decrease depending on the grades students expect to receive. In fact, correlations between the expected grade for a section and the evaluation of an instructor range from 0.10 to 0.30 and average about 0.20 (see Feldman 1997 for review). Interpretations of this correlation include:

1.   Grading Leniency: Instructors that give higher than deserved grades will get higher evaluations. Lenient grading can inflate teacher evaluation scores.

2.   Validity Hypothesis: Better expected grades reflect better learning by students and better teaching by the instructor. Teaching effectiveness causes higher grades and higher evaluation scores.

3.   Students' Characteristics Hypothesis: Preexisting individual differences, such as prior interest in the subject matter, influence learning, grades, and ratings of teaching effectiveness. Any correlation between grades and ratings of teaching effectiveness is spurious.

Studies by Marsh (1983, 1987) and Howard and Maxwell (1980, 1982) show that approximately one third of the relationship between grades and ratings of teaching effectiveness can be attributed to prior subject interest (student's characteristics hypothesis) and two-thirds of the relationship can be attributed to learning (validity hypothesis). They find almost no variance due to grading leniency.

Implementation Issues

Formative versus Summative Evaluation

Centra, in *Reflective Faculty Evaluation*, suggests that formative evaluations will not be as effective in helping teachers if peer reviewers are also making summative judgments, because teachers will not be as open to describing weaknesses or seeking advice from people who will also judge them. In fact, the Stern Business School at NYU adopted a peer review program that

is entirely non-evaluative. The explicit goal of the Stern program is to "raise the level of consciousness about teaching, and make teaching an integral feature of Stern's culture."

The Stern program is designed to increase teaching effectiveness and provides mechanisms for diagnosing and assessing teaching effectiveness. Four alternative diagnostic processes are available and faculty must choose one of the alternatives to use to obtain feedback. Three of the four alternatives use trained consultants. These three methods include (1) videotaping with evaluation by a consultant, (2) audit of a classroom session by a consultant, and (3) discussion with the students in the class by a consultant who then gives feedback to the instructor. The fourth (non-consultant) option is peer review by faculty. The faculty member being reviewed selects the reviewer in consultation with the department Chair. Faculty must report which diagnostic *process* was elected in their annual activity report, however, all reports and diagnostic information (videotapes, etc.) are given to the faculty member at the end of the process and are kept confidential.

Use of a formative evaluation approach, although not consistent with the UF primary objective, has several potential benefits. First, confidentiality sends a clear message that the peer review process is really a mechanism for improving teaching, not an additional evaluation technique. In this way, the process sidesteps the issue of rigorous validity and reliability but is not purely descriptive. This approach also changes the message from a possibly threatening tone, i.e., you do not measure up to the standard, to one of support for real improvement, i.e., we do not doubt your teaching ability, but is there a way to increase your effectiveness. Second, having several options for assessment available to faculty is important as the different processes can be used to assess different aspects of teaching, and different people may feel more comfortable with different types of feedback. Having peer review as an option does allow faculty to get feedback from a subject matter expert, whereas the other three 'consultant' alternatives can only provide feedback on technique.

In order to improve teaching through a formative model, Centra indicates that four conditions must be present:

1.   new knowledge – teachers must first learn something new about their performance
2.   value – teachers must value the information, which means they must have confidence in the source and in the evaluation process
3.   how to change – teachers must understand how to make the changes called for
4.   motivation -- external incentives or internal values must induce change

Criteria, Standards, and Procedures for Peer Review

Centra reports on a review of the literature that identified ten criteria of effective teaching that colleagues are best able to judge. These are

1. mastery of course content
2. selection of course content
3. course organization
4. appropriateness of course objectives

5. appropriateness of course materials (such as readings, media)
6. appropriateness of evaluative devices (such as exams, written assignments reports)
7. appropriateness of methodology used to teach specific content areas
8. commitment to teaching and concern for student learning
9. student achievement based upon performance on exams and projects
10. support of departmental instructional efforts

Centra suggests that percentile ranking and an individual's teaching philosophy statement should be used to evaluate performance on these criteria.   Potential procedures include (1) videotaping class sessions, (2) critiques from trained classroom observers, and (3) use of a teaching mentor. Centra's text includes sample forms for classroom observation and colleague evaluation. Hutchings (1995) also includes various examples of peer review materials and programs that have been used by various colleges.

REFERENCES

Abrami, Phillip C., Cohen, P. A., and Sylvia d' Apollonia (1988), "Implementation Problems in Meta Analysis," Review of Education Research, 58, 151-179.

Centra, J. A. (1979), Determining Faculty Effectiveness. San Francisco: Jossey-Bass.

Centra, J.A. (19xx), Reflective Faculty Evaluation. San Francisco: Jossey-Bass.

d' Apollonia, Sylvia and Phillip C. Abrami (1997), "Navigating Student Ratings of Instruction, American Psychologist, 52 (November), 1198-1208.

Feldman, K. A. (1997), "Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings," In R. P. Perry and J. C. Smart (Eds), Effective Teaching in Higher Education: Research and Practice. New York: Agathon Press, 368-395.

Howard, G. S., Conway, C. G. and S. E. Maxwell (1985), "Construct Validity of Measures of College Teaching Effectiveness," Journal of Educational Psychology, 77, 187-196.

Howard, G. S., and S. E. Maxwell (1980), "The Correlation between Student Satisfaction and Grades: A Case of Mistaken Causation?," Journal of Educational Psychology, 72, 810-820.

Howard, G. S., and S. E. Maxwell (1982), "Do Grades Contaminate Student Evaluations of Instructors?," Research in Higher Education, 16, 175-188.

Hutchings, P., editor (1995), From Idea to Prototype: The Peer Review of Teaching, A Project Workbook. The AAHE Teaching Initiative, American Association for Higher Education

Marsh, Herbert W. (1983), "Multidimensional Ratings of Teaching Effectiveness by Students from Different Academic Settings and Their Relation to Student/Instructor Characteristics," Journal of Educational Psychology, 75, 150-166.

Marsh, Herbert W. (1987), "Student Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future research," International Journal of Educational Research, 11 (3), entire issue.

Marsh, Herbert W. and Lawrence A. Roche (1997), "Making Students' Evaluations of Teaching Effectiveness Effective," American Psychologist, 52 (November), 1187-1197.

Murray, H. G. (1983), "Low Inference Classroom Teaching Behaviors and Student ratings of College Teaching Effectiveness," Journal of Educational Psychology, 75, 138-149.

Peer Review Task Force, Recommendations for Peer Review of Teaching at the University of Florida, Spring 1995.