The Peer Review of Teaching: Progress, Issues and Prospects

Patricia Hutchings

ABSTRACT: As campuses search for ways to raise the level of attention to teaching, the peer review of teaching offers distinct advantages, especially for faculty eager to overcome the isolation of the classroom and to collaborate on improvement. But it presents a number of challenges as well, both political and methodological, and presumes significantly different roles for faculty in ensuring and improving the quality of student learning. Experience on twelve campuses in a national project on the peer review of teaching provides a context for analysis in this introduction to the essays that follow.

Introduction

For several years now, campuses across the country have been reexamining the character of faculty work, looking at the various roles faculty are expected to take on, and particularly considering how the tasks of teaching can be more effectively conducted and improved. One theme evident in resulting reports and recommendations—first appearing in the *Report of the Task Force on Faculty Rewards* ("Pister Report") at the University of California (1991), but now widely heard—is that teaching, like research, should be peer reviewed. In some cases, the call is for peer involvement in high-stakes evaluation of teaching, but there is a general sense, as well, that teaching would benefit from the kinds of collegial exchange and collaboration that faculty seek out as researchers. In what follows, I'll briefly describe current developments related to the peer review of teaching, the rationale behind those developments, issues raised by peer review, and prospects for the future.

Patricia Hutchings directs the Teaching Initiative of the American Association for Higher Education in Washington DC, where she has been a staff member for the past eight years. Prior to that she was a faculty member and chair of the English department at Alverno College in Milwaukee. She received a Ph.D. in English from the University of Iowa in 1978 and continues to teach creative writing.

Current Developments

During the 1970s and 1980s, important strides were made in the evaluation of teaching as student ratings of teacher effectiveness, once the exception, became the rule. Now, in the 1990s, the idea that faculty should be involved as colleagues and peers in the review of teaching constitutes an important step forward in higher education's seriousness about teaching. Evidence of this new stage of evolution can be seen in a number of arenas.

First, there is activity on campuses. In the winter of 1994, twelve universities joined forces in a project to develop new roles for faculty in evaluating and improving the quality of teaching. Each had already made a commitment to rethinking practices related to the evaluation and improvement of teaching, and the project was designed to help similar campuses pursue change together with Northwestern, Stanford, and Syracuse Universities representing private research universities; the University of California at Santa Cruz, the University of Georgia, the University of Michigan, the University of Nebraska at Lincoln, and the University of Wisconsin at Madison representing public research universities; and Indiana University-Purdue University at Indianapolis, Kent State University, Temple University, and the University of North Carolina at Charlotte representing regional and/or metropolitan universities.

Key to the conduct of the project has been pilot work at the departmental level, where faculty roles are shaped and the contexts for attention to teaching are established—or not. Thus, the twelve campuses worked together at the outset to identify a core set of departments, allowing cross-campus collaboration by discipline (e.g., historians at Northwestern working with historians at the University of Wisconsin and Stanford). Faculty teams from pilot departments in eight fields—chemistry, mathematics, English, history, music, business, engineering, and nursing—have now devised and piloted plans for the peer review of teaching in their own settings, with some of their activities now being adapted to additional departmental settings.

The project, entitled "From Idea to Prototype: The Peer Review of Teaching," is funded by The William and Flora Hewlett Foundation and The Pew Charitable Trusts; it is coordinated by the American Association for Higher Education (AAHE), where I serve as project director in partnership with Lee Shulman at Stanford University.[1]

The essays that follow speak to the work of the campuses in the AAHE project, but it is important to say that their stories are but a sample of a larger cloth. Many additional campuses, from a variety of sectors, have recently turned their efforts toward greater faculty involvement in each other's teaching, including its formal review. The California State University system is currently pursuing an initiative on the peer review of teaching and has, on some campuses, well-established, powerful programs of peer coaching and consultation. Campuses ranging from the University of Minnesota-Duluth, to Michigan State University, to George Mason, have recently held their own faculty workshops on the peer review of teaching; and some fifty institutions sent teams to a workshop offered at AAHE's Conference on Faculty Roles and Rewards in January 1995.

Meanwhile—and very importantly—the scholarly societies have begun grappling with new roles for faculty in the review and improvement of teaching. A 1994 report entitled *Recognition and Rewards in the Mathematical Sciences* includes several recommendations related to the peer review of teaching, as does a 1995 report from the Board on Engineering Education entitled *Engineering Education: Designing an Adaptive System*. A spring 1995 Convocation on Undergraduate Science, Math, Engineering, and Technology Education, sponsored by the National Research Council and chaired by former Stanford President Donald Kennedy (who had called for the peer review of teaching on that campus several years earlier) included discussion of the peer review of teaching. The 1995 American Historical Association annual meeting included a session for department chairs focused on the peer review of teaching, and a similar session was offered (to a standing-room-only crowd) at the American Assembly of Collegiate Schools of Business annual meeting in April, 1995. Many of the disciplinary and professional associations are now planning next steps—publications, newsletter articles, conference sessions, and the like—related to the peer review of teaching.

Finally, the peer review of teaching has been a focus of activity outside academe. A recent mandate in North Carolina, for instance, requires all public institutions to conduct peer review, a development

[1]For more information about the AAHE project, "From Idea to Prototype: The Peer Review of Teaching," or to have your name added to the project mailing list, contact: Pam Bender, American Association for Higher Education, One Dupont Circle, Suite 360, Washington, DC 20036.

that Deborah Langsam and Philip Dubois describe in an essay that
follows in this volume. But even in the absence of explicit mandates,
there is a general sense among the public and policy makers that
teaching needs to be higher on the agenda . . . and that the evalu-
ation of teaching, in particular, needs to be more rigorous and com-
prehensive.

In sum, peer review is clearly on the national agenda as an issue
for debate and discussion. There is now a growing body of practice
related to the peer review of teaching, a powerful set of players, a
growing literature on the subject (as evidenced by this volume, for in-
stance) and a sense that this is indeed an idea whose time has come.

The Rationale for the Peer Review of Teaching

As a topic, the peer review of teaching is quick to raise hackles
and arouse fears, especially among faculty. There are good reasons
for such reactions (some of which I'll address in the next section), but
there are also powerful *educational* arguments for the peer review of
teaching, which, in my experience, many faculty find compelling.

First, student evaluations of teaching, though essential and now
widely employed, are not enough; there are substantive aspects of
teaching that only faculty can judge. Currency in the field is an ob-
vious case in point; but, additionally, faculty are best suited to judge
such aspects of teaching as the relevance and power of examples,
what is most important and what can be left out, and what it means
to engage in authentic inquiry in the field. It is important to note
that none of these aspects of teaching is simple or clear cut, and
judgments about all of them must be made in light of context and
purpose. But each is a matter that would benefit from discussion
and judgment by faculty colleagues who know the field and what it
means to understand it deeply.

Second, the peer review of teaching is important because the re-
gard of one's peers is highly valued in academe and teaching will be
considered a worthy scholarly endeavor—one to which large numbers
of faculty will devote time and energy—only when it is reviewed by
peers. Many constituencies—campus administrators, state officials,
federal agencies, foundations, and others—make decisions that shape
faculty lives. But among these constituencies, none is more powerful
than colleagues who are members of one's own scholarly community;
a sociologist at Berkeley looks to many sources for approval and sup-

port—but none more so than other sociologists, both within the department and in other departments around the country. Peer review is the coin of the realm among faculty as professionals; and it follows that, *until* teaching is regarded as worthy of evaluation and review by peers, it will never have the status as a scholarly activity that is prerequisite to widespread commitment from faculty.

Third, the involvement and review of faculty peers is essential because teaching is exceedingly difficult to learn alone. As indicated by recent research on what good teachers know and can do, teaching is a highly complex, situated activity which is learned largely and necessarily through experience. But learning from experience is hard. Watching oneself in action in the classroom, attending to and assessing the multiple aspects of the teaching that go on there—and the learning that does or does not result—is difficult at best. Faculty need each other's help in seeing clearly and making sense of our own practice in order to improve it. And this need is even more pressing in settings where faculty are exploring and attempting to adopt new pedagogies that will be more effective with today's students—moving from concern with covering the material to fostering deeper understandings of it. In short, collaboration and review among faculty peers are essential to educational improvement—and, indeed, something that many faculty would *welcome* as an antidote to the isolation in which teaching has traditionally been practiced.

Finally, peer review is essential because it is a professional responsibility. It is important, certainly, to seek feedback from students about our effectiveness as educators; and in these days of accountability, it is inevitable that other parties, outside academe, will also be passing judgment on the quality of our work. But finally, it is our responsibility as faculty to develop and employ mechanisms for ensuring the quality of our work with students. Though other parties can and should play a role, it is faculty, after all, who can best judge the quality and depth of student learning. Peer review puts faculty in charge of the quality of our work as teachers; as such, it's a right thing to do . . . and a better one, certainly, than having others do it *to* us.

Issues in the Peer Review of Teaching

It is one thing to be persuaded that the peer review of teaching is a right and good idea, but it is quite another to make it happen in ways that will actually be useful. Campuses seeking to establish

peer review practices face a number of hard issues, both political and methodological. An explanation of these issues follows.

1. Going Public with Teaching

Whereas the work faculty do as researchers is public and cosmopolitan, teaching is largely private on most campuses—even those that value teaching highly. "Talking about teaching publicly," writes Diane Gillespie, a faculty member at the University of Nebraska-Omaha, "is like eating somebody else's salad at a formal dinner . . . embarrassing" (1989). And even among faculty who would welcome mechanisms and occasions for sharing what they know and do as teachers, there are legitimate issues related to "going public."

One concern I frequently hear, for instance, is that few faculty have any real pedagogical training or "knowledge base" and thus bring to the review of teaching only their own biases and predilections—a particularly problematic state of affairs when the context is one of high-stakes evaluation of teaching, but a difficulty even where more informal, improvement-oriented exchange is what is wanted. The long tradition of teaching "behind closed doors"—literally and metaphorically—makes any form of peer involvement in teaching a practice that goes deeply against the grain in most academic settings. The good news, as suggested by the essays that follow, is that given the right prompts and occasions many faculty *want* to talk about, discuss, and debate teaching, even in settings where doing so has traditionally been problematic.

2. Establishing Standards

Whether for post-tenure review or the mentoring of junior faculty (i.e. whatever the context of use) peer review presumes some public, shared sense of what constitutes effective teaching. But in fact, on most campuses, there *is* no shared understanding about effectiveness; at best one is likely to find a sort of laissez faire agreement to disagree.

Campuses seeking to move ahead with the peer review of teaching thus find themselves struggling with issues of standards. What is meant by effective (or excellent) teaching? How would we know if it we saw it? What is the relationship between effective teaching and significant student learning? Ought we to expect different levels or kinds of effectiveness from a junior and a senior faculty member? These are not, probably, questions which will or should be finally

resolved; that is, the "standards question" is not one to be taken up, decided, and set to rest. What is needed, rather, is thoughtful, on-going discussion by relevant groups of peers, attempting to say ever more clearly what constitutes good teaching, putting forward power-ful images and examples of it, and working toward a definition that can guide and be guided by concrete acts of judgment.

It may be, too, that what is most useful is not a definition of good teaching but a clearer, more public agreement about what constitutes significant student learning. As many of the faculty in the AAHE project have noted (and as proponents of student assessment have argued now for a decade) the test of teaching is, after all, student learning. There are difficulties in this formula in that learning de-pends on what students bring to the equation as well as what teach-ers do; but at the very least, one aspect of peer review of teaching that needs further work is attention to the effects of teaching on student learning as part of the larger picture of teaching excellence.

3. Identifying the Appropriate Peers

A common question about the peer review of teaching pertains to the identity of the peers. "Who's to judge?" one must ask. Is the rele-vant observer/judge a colleague from one's own department? Someone from another field? A colleague from another campus?

Some, like Lee Shulman, weigh in on the side of disciplinary peers. "We would never dream," Shulman remarks, "of sending out exam-ples of someone's research to people at another university who *in general* were on that other university's faculty. The medievalists get the research reports of other medievalists, and the civil engineers receive the reports of other civil engineers." Thus, he concludes, we must "make the review and examination and support of teaching part of the disciplinary community's responsibility . . ." (Shulman, 1993).

Uri Treisman, whose work in mathematics education won him the prestigious Dana Award, puts the issue somewhat differently, begin-ning not with questions of the discipline per se but with questions about the relevant problem-solving community. Treisman's opinions are articulated in a report entitled *Teaching growth and effectiveness: An issues paper* (1994). Treisman would say, for teachers working on calculus reform, the relevant peers are other teachers of calculus around the country working on that significant educational problem. Following this thinking, for the faculty member who wants to docu-ment her ability to help students connect ideas across disciplines,

the relevant problem-solving community might be other faculty on campus who are invested in the design and teaching of the general education curriculum.

The punchline here is that the "community of judgment" for teaching is actually several different communities. Some aspects of teaching, like classroom performance, are probably best reviewed by peers within the institution; others, including syllabi and other materials, might benefit from external review by peers at the national level. Some aspects require peers in the field; others benefit from the perspective of faculty from different areas of study. What's needed now is more extensive experience with the usefulness for various purposes of one set of peers versus another.

4. Finding the Right Methods and Strategies

Even in settings where there is general consensus that peer review is a necessary next step, there is little clarity or agreement about what methods to employ. Though extensive literature exists on selected issues of method (French-Lazovik, 1976; Cohen & McKeachie, 1980; Cross, 1987; Weimer, 1990; Kahn, 1993), it remains an open question what documents and evidence—in what combination—are most revealing of what teachers know and do. Further, one must ask how the answer to this question may vary by context. Are the key artifacts in the peer review of biology teaching different from those for freshman composition, for instance? How do the "rules of evidence" in various fields shape what will and will not be credible and useful by way of documentation.

To put the issues of method and strategy more concretely, one might ask whether it is more useful, for example, to review a course syllabus or, another possibility, actually to observe a classroom interaction. If the latter is too intrusive, as many faculty insist, what about videotape? And if video is promising, should the focus be on a single class session? More than one? What additional information do you get with two cameras rather than only one? Or, to bring the question full circle, are videotapes useful only when combined with and contextualized through examination of syllabi? . . . .

A related issue has to do with exactly how evidence about teaching is communicated—in what genres or forms. Some faculty are exploring the use of narrative, or cases, to convey the wisdom of practice, for instance. Those who have embraced teaching portfolios, on the other hand, put the focus on "work samples" (a syllabus, for example, or

examples of student performance), combining these with reflective commentary wherein the teacher might explain how the syllabus has changed over time or provide background about the student population.

"Hard data" and numbers are what others look to and find compelling. But the relative usefulness of these and other types of evidence needs further exploration, especially as related to audience. What is useful to colleagues in the department may be quite otherwise to a cross-disciplinary faculty group trying to make decisions about a teaching award or about promotion and tenure. And the character of evidence that might speak convincingly to outside audiences, be they legislators or prospective students, is yet another matter.

Finally, there is the unit-of-analysis question. Many of the usual sources of evidence about teaching effectiveness assume that the important level of analysis is the single class incident or episode, e.g. a classroom visit by a colleague or a single piece of student work. But professor William Cerbin (1994) at the University of Wisconsin-La Crosse has argued for and developed a prototype of a *course* portfolio—where the object of potential review is not individual class sessions/assignments/episodes but the larger act of conceiving and directing a course of study. A number of faculty in the AAHE project have also begun developing course portfolios. Working on an analogy of teaching to a scholarly project, their portfolios look at 1) goals and intentions (usually in the form of a syllabus or course design), 2) enactment (usually through materials and evidence related to what actually goes on in the classroom), and 3) results (most importantly through evidence of student learning). Portfolios so constructed capture the intellectual integrity of teaching in a way that many strategies do not. Additionally, the process of developing the portfolio, as faculty who have developed them report, is a powerful prompt to reflection about assumptions behind the conduct of the course and about ways it might be altered and improved (AAHE, 1995, tab 4).

For many people, the peer review of teaching is synonymous with classroom observation; but what is needed is a range of strategies and multiple sources of evidence, that, in proper combination, provide a more comprehensive picture of the various aspects of teaching, in ways appropriate to the range of purposes for which peer review is employed.

5. Time

In a recent survey (May 1995) of campus participants in the AAHE project on peer review, the majority of faculty respondents said that

the greatest obstacle to progress was insufficient time. They and
their departments valued teaching, and they believed in the impor-
tance of greater peer involvement in ensuring and improving the
quality of teaching; but "not enough time" had kept them from en-
acting their peer review plans as fully as hoped. This, it need hardly
be said, is an issue affecting virtually all campuses, especially as the
roles that faculty are expected to perform have proliferated in recent
years.

  One implication is that our approaches to peer review must be
designed to maximize benefits and take advantage of "spinoff effects."
The course portfolio, for instance, as explained above, is useful not
only in its ability to present a more intellectually coherent picture
of teaching but in its power to prompt reflection and improvement.
A second implication is that campuses and departments experiment-
ing with peer review must pay close attention to costs and benefits
and be prepared to share what they learn with others.

  But beyond these practical imperatives, the issue of time brings
us to questions of values, of what tasks we believe are most worth
doing. Faculty on many campuses take for granted the need to spend
large amounts of time documenting their own research and peer re-
viewing the research done by others; in contrast, the amount of time
spent documenting and peer reviewing teaching is minuscule. "I am
not impressed when colleagues who happily edit or review for jour-
nals and grant agencies whine about the time it might take to look
at teaching," one participant in the AAHE project told me. "I know
they're busy; I'm in the same boat. The issue is one of priorities, of
giving greater professional time and thought to teaching, even if at
the expense of something else" (Bernstein, 1995).

Prospects for the Future

  Establishing meaningful processes for the peer review of teaching
will be a long-term effort, one with implications for fundamental as-
pects of academic culture and values. As one participant in the AAHE
project noted, "We're talking about changes that I don't expect to see
in my lifetime."

  Nevertheless, a look at efforts currently underway suggests a
number of themes—not quite "how-to" tips, more like shaping prin-
ciples or assumptions—that would seem to characterize promising

work. Here are three that seem especially important to future work of the AAHE project and related efforts.

A Vision of Student Learning

The single most important reason to move toward the peer review of teaching is the need for more powerful student learning. As faculty in many settings are realizing, the goal of college and university teaching in the 21st century must be *understanding;* the reigning emphasis on facts, and on mastering information, should give way to more active forms of learning—forms that bring students to deep understanding and engage them in making meaning. Progress on this difficult front means attention to the kinds of teaching that engage students more deeply and thoughtfully in subject-matter learning and in making connections between their lives and their academic studies; it means turning classrooms into communities of scholarly inquiry in which students can be authentic participants.

The radical nature of this transformation—for faculty and students both—cannot be overemphasized. It is a transformation that will require from all of us new behaviors and ways of thinking about what we do.

A New Conception of Teaching as Scholarly Work

If deeper understanding by students is the goal, we must begin to think about and treat teaching as more than technique. As Ernest Boyer argued in *Scholarship Reconsidered* (1990), teaching is an aspect of scholarly work; choices about course design, assignments given, criteria for evaluating student learning—all are reflections of the way the teacher thinks about his or her field, and what it means to know it deeply. Unfortunately, these more substantive, intellectual aspects of teaching have often been ignored or overlooked in the attention to technique and method that typically prevails in discussions of teaching. A central contribution of peer review can be to capture the scholarly aspects of teaching that demand to be talked about if we are serious about improving student understanding. Or, to put it differently, peer review will be distinctly powerful where it is aimed at capturing the scholarly aspects of teaching.

This lesson was driven home in the AAHE project when faculty were asked to do three "exercises" focused, respectively, on the *syllabus* as the plan or prospectus, *classroom practice* as an enactment of that plan, and *student learning* as evidence of results. Each of the

exercises called for an artifact plus reflective memo (*From Idea to Prototype*, 1995, tab 3).

Faculty responses to these three exercises deserve more in-depth analysis than can be tackled here, but the most important point is that the exercises prompted wonderful, thoughtful accounts of teaching—accounts in which teaching is treated not as disembodied, generic technique but as the *representation and transformation of ideas.* For example, an English professor from Indiana University-Purdue University at Indianapolis explored the implications for a senior capstone course of the fact that "the discipline of English Studies is constantly redefining itself;" a member of the chemistry department team from the University of Wisconsin described the rationale behind "a cooperative group project on determining the amount of carbonate in a sodium hydroxide titrant using the H3PO4 system." In short, the exercises prompted a kind of discourse about teaching that is unusual, in my experience, and that clearly captured faculty energy and attention. At a project meeting organized around the exercises, most everyone seemed to want more time to discuss their responses with others, especially others in their own field.

This is not to suggest that cross-disciplinary talk about teaching is not powerful stuff; rather, it points to the need for a kind of discussion and exchange that often *doesn't* occur on campuses, where teaching improvement efforts typically are organized across disciplines and focused on technique. What is needed is discussion of teaching as the transformation of ideas for student understanding and that's where peer review can make its mark: an occasion for faculty to talk about the scholarly, intellectual work of teaching "my subject to my students."

New Roles for Faculty in Ensuring and Improving the Quality of Teaching

To capture the scholarly substance of teaching requires more active roles by faculty in assembling the picture of what they do and in revealing the thinking behind the choices they make. As things now stand on most campuses, faculty are not actively involved in documenting what they do as teachers. The evaluation of teaching, for instance, seems almost to happen *to* faculty, as objects rather than as active agents in the process. But it is not enough to depend on student views to represent teaching; nor is it sufficient to drop a syllabus into a promotion and tenure file. The need is for faculty to

be more active agents in putting together appropriate artifacts of their teaching, along with reflective commentary that reveals the pedagogical reasoning behind them.

As one of the faculty participants in the AAHE project noted, no one, prior to the project, had ever asked him to explain the thinking behind his work as a teacher; and to do so put him in a different and very welcome role, not as the object of someone else's critical scrutiny but as a scholar sharing his thinking with others who have similar interests.

In this sense, the peer review of teaching can, in its most powerful forms, be less a matter of *judging teachers* than of *improving teaching,* with the focus moving increasingly to ways we can help each other improve the quality of our collective contribution to students' learning. That would be a change, indeed—one with the potential to make a real difference where it matters most.

For more information about the AAHE project, "From idea to prototype: The Peer Review of Teaching," or to have your name added to the project mailing list, contact: Pam Bender, American Association for Higher Education, One Dupont Circle, Suite 360, Washington, DC 20036.

Acknowledgments

References

Bernstein, D. J. (1995, September). Personal communication.

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate.* Princeton, NJ. The Carnegie Foundation for the Advancement of Teaching.

Cerbin, W. (1992). How to improve teaching with learner centered evaluation. *National Teaching and Learning Forum, 1*(6), 6-8.

Cohen, P. A. & McKeachie, W. J. (1980). The role of colleagues in the evaluation of college teaching. *Improving College and University Teaching, 28*(4), 147-154.

Cross, K. P. (1987). *Feedback in the classroom: Making assessment matter.* Washington, DC: American Association for Higher Education.

Engineering education: Designing an adaptive system. (1995, April). Report of the Board on Engineering Education (pre-publication draft), National Research Council.

French-Lazovik, G. (1976). *Evaluation of college teaching: Guidelines for summative and formative procedures.* Washington, DC: Association of American Colleges.

From idea to prototype: The peer review of teaching (a project workbook). (1995). Washington, DC: American Association for Higher Education.

Gillespie, D. (1989, July/August). Claiming ourselves as teachers. *Change, 21*, 56-58.

Kahn, S. (1993). Better teaching through better evaluation: A guide for faculty and institutions. *To Improve the Academy, 12*, pp. 111-126. Stillwater, OK: New Forums Press.

Recognition and rewards in the mathematical sciences. (1994). A report of the Joint Policy Board for Mathematics Committee on Professional Recognition and Rewards.

Report of the task force on faculty rewards. (1990). Report of the University of California System.

Shulman, L. S. (1993, November/December). Teaching as community property. *Change, 25*, 6-7.

Teaching growth and effectiveness: An issues paper. (1994). Report of the Mathematical Sciences Education Board, National Research Council.

Weimer, M. (1990). *Improving College Teaching*. San Francisco: Jossey-Bass.