Peer Review of Teaching: Lessons Learned from the University of Kentucky

Victoria M. Bhavsar,* Larry J. Grabau, and M. Scott Smith

ABSTRACT

The University of Kentucky Department of Agronomy instituted a system of peer review of teaching in 1996. Professors observe the teaching of their peers with the intent of helping one another improve. Reviews usually include a preliminary meeting between the instructor and two reviewers, several in-class visits, some analysis of classroom materials, and a wrap-up meeting. The peer review program now has met with cautious but general approval from participating faculty. By spring 1999, reviews were completed for 20 classes, 15 of which are considered in this analysis. Overall departmental student ratings have not changed as a result of peer review, but most participants have made changes in their teaching as a direct result of review. Although instructors are not required to place peer review data in their personnel dossiers, six reviewed professors included data in their 1998 files. Although peer review has not yet made substantial differences in the way that teaching is rewarded in the department, most professors consider peer review professionally valuable and personally satisfying. Our experience indicates that the most successful reviews require the instructor and reviewers to have open attitudes, spend adequate but not excessive time on reviews, and receive support from administrators and consultants. Peer review is effective in helping this department to open teaching to discussion and recognition.

The past president of the American Association for Higher Education says, “In my mind, [peer review of teaching] is the single most promising trend of all [for the improvement of teaching]” (Edgerton, 1993). Peer review has gained attention, partly because many faculty members want teaching to be accorded the prestige and reward that research enjoys, which require that teaching, like research, be openly discussed and scrutinized. Careful peer review of teaching can foster such activities. Peer review has also gained attention because institutions are pressured to account for the quantity and quality of faculty work in the classroom (Seldin, 1993).

Seldin (1993) surveyed research institutions, observing a 10-yr trend toward more formal, less haphazard data gathering to evaluate teaching. In 1998, Seldin performed a similar survey of liberal arts institutions and found a 10-yr trend toward using diverse data sources to evaluate teaching. Although teaching evaluation is generally improving, some recent papers still report that teaching is often evaluated with superficial information like hearsay (Massy et al., 1994; Way, 1992). Mundy et al. (1999) found that teaching evaluation in plant and soil science and agricultural economics departments varies in effectiveness and that student ratings are by far the most prevalent form of teaching evaluation in these departments.

The faculty of the University of Kentucky (UK) Department of Agronomy designed and implemented a system of peer review for teaching. The overall goal is to encourage a departmental culture in which teachers are well prepared and highly valued. The program’s first objective is to recognize and spread excellent teaching strategies among faculty. The second objective is fair, thorough evaluation of teaching efforts at the departmental level. This paper describes our program’s history, current policies and procedures, and lessons learned from 3 yr of peer review.

Overview of University of Kentucky Agronomy

Agronomy is the largest department in the UK College of Agriculture. In 1998, agronomy had 45 full-time faculty with 29 full professors, 10 associate professors, and 6 assistant professors. In the academic years 1996 to 1999, 2 professors had teaching appointments of >50% of their distribution of effort; 5 professors had 35 to 50% teaching appointments; 11 had 11 to 34%, and 8 had 10% or less. Nineteen professors had no official teaching appointment but might teach or team-teach one class every other year. Total full-time equivalents (FTE) in the department included 6.4 FTE teaching, 24.2 FTE research, and 14.4 FTE extension in 1996 to 1999. Conversations with about 30 professors indicate that individual faculty members’ interest in teaching varies from great personal devotion and professional commitment to indifference, with more than 50% liking teaching but placing more emphasis on other activities.

Agronomy faculty have averaged 2830 total student contact hours per year since 1990–1991 by teaching courses in three curricula: plant and soil science, agricultural biotechnology, and natural resource conservation and management. Most College of Agriculture students take some courses in these programs, so the probability of nonagronomy majors having at least one professor from agronomy is high. The quality of teaching by agronomy professors is important to education in the College of Agriculture.

History of Peer Review in University of Kentucky Agronomy

In 1993, the department head raised the idea of mandatory peer review of teaching because he was concerned about his over-reliance on the mandatory student rating forms for his evaluation of teaching. Two faculty members with large teaching appointments were very concerned about the quality of teaching in the department and seconded the chair’s idea. At that time, more than half the faculty voted against peer review in an anonymous paper survey. Faculty said they did not have time for in-depth review of one another’s teaching; neither did

Abbreviations: UK, University of Kentucky; FTE, full-time equivalents.

Department of Agronomy, N122 Agricultural Sciences North, Univ. of Kentucky, Lexington, KY 40546-0091. Received 6 Aug. 1999. *Corresponding author (toria@pop.uky.edu).

http://www.JNRLSE.org

they want to open their teaching to public scrutiny when both the process and implications of doing so were unclear.

In 1995, the department head raised the idea of mandatory peer review again. This time the faculty accepted peer review almost unanimously. They were encouraged by an influential senior professor’s interest in having his own class reviewed, and by the endorsement of the department’s undergraduate education coordinator, a respected faculty member. A proposal to the USDA Higher Education Challenge Grant Program requested funds for a staff member to coordinate reviews. The associate dean for instruction in the College of Agriculture supported the idea of peer review, suggesting that a model could be presented to other departments if agronomy succeeded. The campus Teaching and Learning Center staff offered ideas and encouragement as a departmental committee wrote the first draft of a policy for peer review of teaching. The committee used Keig and Waggoner’s (1994) publication Collaborative Peer Review as a focal point for discussion.

As plans progressed in 1995, the faculty disagreed about how to use review data. The department head wanted to use review reports for personnel evaluation (summative evaluation). He also believed that including review reports in files was the only way to give credit for participation. Many faculty members objected, concerned that they could not give honest feedback if the information were used for personnel decisions. They preferred to use review information only for self-improvement (formative evaluation). Ten professors believed that the system should be flexible enough to accommodate both summative and formative purposes. The issue remained unresolved throughout 1996.

In 1996, four peer reviews were performed for established professors who had already participated in reviews; useful feedback was limited. All professors involved reported dissatisfaction but wanted to give the project time to develop. Other faculty members expressed doubts about the value of peer review.

Late in 1996, the proposal to USDA was funded and the department hired a coordinator in January 1997 to organize the peer review effort. She held a professional staff position in the department with responsibility to organize peer reviews, evaluate the program’s success, and act as a general teaching resource person. She organized a small training workshop in February. She did not serve as a reviewer but offered help with scheduling, ideas for review methods, and other support. Three courses were reviewed in the spring 1997 semester.

A larger training workshop occurred in the summer of 1997. Approximately 22 faculty members attended, including most professors who had already participated in reviews. They practiced with various review techniques. A lively discussion concluded with agreement that review results were not required in personnel files but could be included at the reviewer professor’s discretion. The department chair was disappointed with this outcome; he yielded since the faculty would accept peer review only under this condition.

Four more courses were reviewed in Autumn 1997. Faculty expressed cautious, general satisfaction with peer review. In 1998, three courses were reviewed in spring semester and four were reviewed in the autumn. In 1999, two were reviewed in spring, for a total of 20 classes reviewed since 1996.

Current Policy of Peer Review in University of Kentucky Agronomy

Each teaching faculty member is peer reviewed at least once every 4 yr. Professors may request review at any time. The department head selects courses for review, usually scheduling three to four reviews per semester. All faculty members except the department head may serve as peer reviewers, even nonteaching faculty. A professor being reviewed nominates four potential reviewers. The department head appoints two of the four nominees to be reviewers. No faculty member is required to review more than one course per year. To avoid overburdening reviewers and to encourage wide participation, no one is allowed to review classes more than two out of three semesters.

The instructor and his or her reviewers hold a prereview meeting to discuss course goals and methods, and work out details of the review process. Reviewers must make at least three in-class visits, except for classes where visiting is clearly inappropriate, such as modular courses in which more than five instructors give short lecture series. They should scrutinize materials such as the course syllabus, assignments, and tests, and supplementary learning materials. The instructor decides whether to use special activities such as videotaping or student interviews, and whether to present information from such activities to reviewers.

Reviewers should provide written and/or oral feedback that identifies strengths and weaknesses and gives clear recommendations for improvement. The instructor should describe specific improvement steps. The group is encouraged to hold a wrap-up meeting to summarize, discuss recommendations, and produce a written summary. Reviewed faculty may include summaries in their professional files if they feel that summaries would improve administrative evaluation of their teaching.

RESULTS OF PEER REVIEW

The review coordinator interviewed the department head and faculty members who have served as reviewers or been reviewed since 1996. Some people were interviewed more than once since they participated in multiple reviews; interview results may reflect the opinions of these people more heavily. Interview notes were analyzed by pattern coding (Yin, 1997). Collective results from 15 of the 20 reviews, excluding pilots and one ongoing review, are presented.

Process of Peer Review

Face-to-face prereview meetings occurred in 14 of the 15 reviews. These meetings usually lasted less than 1 h, long enough for the teacher to provide a syllabus and course materials, identify dates for reviewers to visit classes, and touch on course goals and methods. The peer review coordinator assisted several teams in scheduling and planning premeetings.

In five reviews, reviewers visited a total of three class meetings (i.e., each reviewer went once alone, then both visited once together). In six reviews, two reviewers visited twice together. Peer review literature generally recommends
at least two visitors visiting three times each for reliability and validity (Braskamp, 1994). Although the peer review policy states that reviewers must visit three times each, the coordinator and department head did not force the issue, preferring that reviewers visited willingly even if they visited fewer times than they should.

In two reviews, the teams made four visits, not remembering if they visited alone or together. One energetic reviewer visited a class 15 times. For another class, one reviewer visited five times and the second reviewer visited nine times. Reviewers who visited this frequently reported that they learned a lot since they saw many different teaching techniques demonstrated. Frequent visitors also felt that their assessments of classroom situations were much more valid because they saw a broad array of events.

Some reviewers visited three or four consecutive classes, while others visited classes randomly during the semester. This choice depended on whether the teacher was more interested in having reviewers comment on continuity from class to class or on general teaching quality. One common concern with classroom visiting is that the teacher will act that day, giving a better-than-usual performance because peers are present. Only one teacher admitted to doing this; all other professors said that they tried not to behave differently and that the presence of reviewers was not overly distracting for them or their students. All professors told their students about the peer review.

Conversations between reviewers and teachers the same day as class visits occurred in nine reviews. Often only one reviewer participated in such immediate feedback. If reflective conversations did not occur within the day, they often did not occur at all unless a postreview meeting was planned. Professors who received immediate feedback greatly appreciated it, even if conversations lasted only a few minutes.

As part of their overall reviews, seven instructors requested midsemester student interviews. The peer review coordinator performed the interviews, using the Small Group Instructional Diagnosis, a method that elicits comments from the entire class in one period (Diamond, 1988). These instructors all valued information from student interviews, three even saying that student feedback was the most useful part of the review. Teachers could integrate student feedback into peer review by asking reviewers to comment on student concerns.

Reviewing materials such as textbooks, exams, and homework proved much more difficult and time-consuming than visiting classes. Reviewers felt that they could not adequately judge materials without extensive class visiting. For example, connections between tests and lectures could not be assessed just from reading test questions. Only three reviews had substantial time given to materials not mentioned in visited lectures. In five reviews, some attention to materials was reported, while in seven reviews minimal time was reportedly given to materials. All professors provided reviewers with at least a syllabus. Two or three reviewers analyzed syllabi carefully, but most reviewers did not comment on syllabi unless inadequacies were obvious. Because of the lack of attention to materials, reviews have tended to be one-dimensional. One reviewer commented, “I reviewed the lectures, not the [whole] class.”

Six instructors reported that they included peer review reports of some kind in their 1998 personnel files. The department head reported that “one or two” teachers forwarded substantive output to him, and “one or two” teachers included typed summaries in their teaching portfolios. Other instructors and reviewers merely wrote lines in their vitae to report their participation.

**Reported Benefits of Peer Review**

While quantitative measures of improvement such as student rating numbers have not changed thus far, 18 of 22 professors who participated as teachers or reviewers reported that they made changes in their teaching and/or extension presentations because of reviews. One teacher even remarked, “The reviews have helped me improve my teaching about 40%. If not for the reviews, I’d be terminally pessimistic about teaching, but [now] I have a glimmer of hope.”

At least six teachers made changes in visuals and handouts because of comments in peer reviews. Two or three reviewers even modified their own materials. Although such changes seem small, one reviewed professor commented on their importance: “I want students to concentrate all their effort on learning and none on overcoming unnecessary obstacles like overheads with small print.”

Peer review has also helped some professors become more aware of significant events in the classroom. Two professors who presented detailed, fast-paced, unorganized lectures began to use outlines and to encourage student participation and presentation because their reviewers observed that the students were lost in the flood of information. Another professor realized that his brusque manner intimidated students and consciously tried to be more gentle; three students reported informally that they became much more comfortable with participating.

New teaching techniques have spread from one teacher to others through peer review and through conversations begun by peer review. For example, a new faculty member had an innovative grading scheme that a senior teacher modified and adopted for his own class; the plan has piqued interest among several other teachers. One professor immediately adopted a suggestion about how to encourage students to read material before classes; his reviewers will probably also adopt it.

Only two participants reported that they had ever visited each other’s classes before peer review began; other participants had probably never before visited their colleagues’ classes or looked at syllabi. Among other factors, peer review helped to bring classroom data to departmental curriculum discussions. One professor who taught upper-level soils classes was extremely interested to discover what students learned in the introductory soils class.

Peer review seemed to benefit reviewers by giving them quick refreshers on old material and sometimes introducing them to new material. Several reviewers remarked that they were as interested in the material presented as in the teaching techniques, and one person even admitted that he was so fascinated by the content that he forgot to review the teacher. Four reviewers attended class field trips, which they described as “a lot of fun” and “a good break.” Reviewing someone else’s class gave faculty members time to pause briefly in their own projects while learning about other work in the department.

In several cases peer review encouraged collegial relationships. At least two professors reported that spending time
with colleagues they would normally not see much was pleasant and stimulating. One reviewer remarked that it was personally satisfying to help someone he regarded highly. Another person commented that reviewing was “not a hoop to jump through,” but that “talking with colleagues about teaching was affirming and heartening.”

Another tangible benefit of peer review is that a core of excellent teachers in the department has become publicly recognized. The good reputations of several professors were confirmed by their colleagues. One reviewer commented, “It’s reassuring to know first hand that I’m sending my students to a good teacher and good subject matter.” The department head agrees that the increased talk about teaching has led to everyone knowing the best teachers and according them added respect. Peer review has begun to achieve our program goals as well as beginning to change teaching from a private endeavor to, in Shulman’s (1993) words, “community property” in our department.

**Downfalls of Peer Review**

While most faculty comments about peer review have been positive, not all the results have been good. In two reviews, participants reported feeling hurt, ignored, or even angered. In one case, the teacher and one of his reviewers had had past conflicts and should not have been matched in a review. Conflict due to a misunderstanding terminated that review process. In the second case, the two teachers of a team-taught class each had many more years of experience than their reviewers. The reviewers felt dismissed and ridiculed, and they worried that serious critique of their senior colleagues would “come back to haunt them.”

Two reviews seemed neutral, with little positive or negative feedback given because of the disengagement of reviewers and/or teachers. One reviewer shrugged and said, “[The instructor] was doing a good job. I didn’t see a way he could improve, so I thought it wasn’t much use. Maybe if you went to a guy who was doing a bad job [it would be useful].”

The other reviewer for this class did provide some constructive suggestions, as did student interviews. The teacher in the second neutral review felt that the reviewers concentrated too much on style suggestions, and his 30 yr experience allowed abundant time to polish style. His reviewers, though, felt that this teacher would not accept comments on anything but surface issues.

The hardest part of the reviews for every participant was writing reports or summaries. When plenty of informal feedback had taken place, reports seemed redundant. Some reviewers were nervous about committing critique to paper or felt that their comments would be unread by uninterested teachers. At least six reviewers simply did not take time to write until long after the semesters were over, if then. In two reviews the teachers received such minimal comments that no change to their teaching could be reasonably expected.

Lack of reporting was a problem for the department head as well as the teachers. Even though he had agreed to formative use of review data, he was troubled that he received limited, qualitative review information inappropriate for merit decisions. More positively, he perceived that instructors who included reports in personnel files or teaching portfolios were committed to improving. He believed that faculty members who were asked to serve as reviewers had the respect of their peers. These perceptions were especially important in his evaluation of young professors. The associate dean of instruction in the College of Agriculture approved of peer review although it probably made no difference in his ratings of faculty merit. A problem with administrators’ current perception and use of peer review information is that minimal review efforts could count as much as serious efforts. Faculty could mitigate this problem by reporting their reviews in some detail in their teaching portfolios.

**LESSONS FOR SUCCESS**

The most important lesson we learned is that attitudes toward review are the primary factor in determining the success of each peer review event and the program as a whole. When professors want to learn from one another, reviews will be useful, even if only a few hours are spent. One instructor said, “I expected it to be extremely helpful and it was.” Positive attitudes can spread if people talk to one another about their experiences. Three professors have actually requested review, while others are now more willing to participate when asked than when we began.

The coordinator was necessary, at least at first. We found that one person must act as organizer so that review logistics are arranged smoothly. Once the program was established, administration time dropped to 1 or 2 h/wk. We hired a special full-time coordinator, but a faculty member released from some other responsibility could do the task. If several departments became interested in peer review or other teaching programs, they could consider ways to hire a specialist who could serve multiple departments.

The department head’s support was essential. His commitment to peer review for personal faculty development was evident when he relinquished his desire to require reporting of review information so that faculty would be willing to participate. Three or four professors felt the program would have failed if the head had not promoted peer review by letting go of its summative aspect, by asking people to participate, and by giving support to the coordinator.

We learned that peer review information could not support personnel decisions. Instead, administrators needed tools such as teaching portfolios to gain well-rounded views of professors’ efforts. Our faculty’s refusal to participate in summative peer review is supported by faculty development literature (Weimer et al., 1988) and in the experiences of other departments, which have found that asking peers to grade one another’s teaching leads to bland, uninformative reviews (Mundy et al., 1999). Also, we found that the faculty needed to agree how review information would be used before they participated in review.

We learned that review processes should be tailored for the needs of each class or instructor and not be mandated in a policy, although some procedures such as premeetings, student interviews, and frequent class visiting were often useful. Training helped teams to learn different review methods and create substantive pedagogical questions. All faculty participating in reviews received notebooks with examples of questions, checklists, review forms, and articles, but few people used the materials unless they attended the summer workshop.
More training events besides our two workshops would have helped to maintain progress, but faculty declined more events. We found that the review program need not and should not be limited to the teaching clique. Even people deeply committed to teaching and personal development grew tired of peer review; the effort needed to be spread out. Second, although extension education is different from undergraduate teaching, extension faculty reported that they could contribute to and receive benefit from reviewing classroom teaching. Third, research faculty also benefited from the change of pace and new ideas provided by reviewing.

Choosing reviewers is a many-sided issue. Most importantly, reviewers and teachers need to trust and respect each other before reviews begin. Reviews may worsen conflictual relationships, and strengthen good relationships. We found that all but two or three instructors carefully tried to choose reviewers who would provide useful, constructive criticism; there was no tendency for instructors to choose buddies. The respective disciplines of reviewers and teachers was not critical unless one of the teacher’s goals for review was to change course content substantially or unless special pedagogical problems were directly related to the discipline. We found that having one disciplinary reviewer and one reviewer outside the discipline (e.g., a soil chemist and a plant breeder for an introductory soils course) could provide a balanced perspective.

Another important issue in choosing reviewers was the willingness of potential reviewers to devote some time. A useful, thorough review required a reviewer to visit three or four classes over the semester, to pay attention while visiting, to spend some time after each visit with the teacher, then possibly to spend an hour writing a report. Review of classroom material required 3 to 4 h during the semester. Reviews could take as much time as a reviewer wanted, but satisfactory reviews did not need more than about 10 h over the course of a 16-wk semester.

Reviews that were heavily based on classroom visits required considerably more time during the semester from reviewers than from teachers, although a teacher could take a great deal of time later on to respond to suggestions. Visiting and informal feedback was much less burdensome than reporting. Immediate informal feedback provided the most information to teachers for the least time invested by reviewers. We now suggest that review teams minimize reporting and replace it with immediate informal feedback and a brief postsemester summary meeting. When review teams hold premeetings to schedule visits to classes, they should also identify half-hours to talk within the day or two after a visit. These talks relieve pressure on reviewers to write detailed reports, and teachers can learn more from discussing a class while it is fresh in their minds. Another important outcome of immediate feedback is that teachers can use advice within the semester to the benefit of current students.

CONCLUSION

Faculty benefit from participating in peer review by improving their own teaching, helping colleagues, strengthening relationships, and getting glimpses of the range of work in the department. The most important result of our 3-yr program has been a small but identifiable change in teaching. The success at UK has not come without effort—people have spent time on reviews, strong disagreements over process and the use of peer review reports have been resolved, and some relationships have had to be mended.

Professors are more aware of classroom events and their own actions and have responded by making changes to help students learn better. Conversations about teaching have become more frequent in the department, and effective teaching strategies are spreading. A recognizable group of excellent and committed teachers and reviewers has become visible and gained respect in the department. Peer review has been a useful tool in helping us to encourage teaching to become community property.

REFERENCES


Other Useful References