**Peer Review of Teaching***by Joe Bandy*

**Introduction**



|  |  |
| --- | --- |
|  | In higher education, peer review stands as the prime means for ensuring that scholarship is of the highest quality, and from it flows consequential assessments that shape careers, disciplines, and entire institutions.  While peer review is well established as a means of evaluating research across the disciplines, it is less common in the assessment of teaching.  Yet it is no less useful, since it can improve what Ernest Boyer has called the “scholarship of teaching and learning” by enhancing instructional and faculty development, by bolstering the integrity of personnel decisions, and by enabling more intentional and mutually supportive communities of scholar teachers.  This guide is intended as an introduction to the basics of peer review, including its purposes, challenges, and common practices.  The primary audience for this guide consists of departments, programs, or schools considering implementing peer review, although individual faculty, staff, and students are likely to find what follows interesting, as well. |

**What Is Peer Review?**

Peer review is often identified with peer observations, but it is more broadly a method of assessing a portfolio of information about the teaching of an instructor under review.  This portfolio typically includes curricula vitae, student evaluations, self-evaluative statements, peer observations, and other evidence such as syllabi, assignments, student work, and letters solicited from former students.  This said, peer observations will figure prominently in what follows.

It is also worth noting a common distinction between two very different forms of peer review: formative and summative.  *Formative evaluation* typically is oriented solely towards the improvement of teaching and is part of instructional mentorship and development.  *Summative evaluation*, in contrast, is that done to inform personnel decisions.  To improve the freedom and exploration of individual faculty, formative reviews may be shielded from scrutiny for a period of years until such time that there needs to be accountability to standards of excellence for personnel decisions.  At this point in time, summative evaluations are more common since they are tied to decisions related to reappointment, promotion, or tenure (Bernstein et al. 2000).  Because the more consequential nature of summative evaluations tends to diminish the formative value of the peer review process, it is important to maintain a clear distinction between these types of evaluation and be transparent with those under review.  It is also common to have different faculty involved in each form of assessment – mentor faculty in the formative evaluation and departmental or program administrators, such as chairs, involved in summative evaluations.

**Why Peer Review?**

Peer review serves many functions in the process of evaluating faculty, courses, or entire programs.

**What’s good for research is good for teaching.** As in peer reviews of research, it is a vital means of receiving expert assessments of one important part of scholarly practice: teaching.  As with research, peer review ensures that faculty internalize, in the words of Pat Hutchings, scholarly “habits of mind” by identifying goals, posing questions for inquiry, exploring alternatives, taking appropriate risks, and assessing the outcomes with learned colleagues.  When this process of scholarly engagement and deliberate improvement is part of the institutional expectations for teaching, as it is with research, it can function to support a community of scholarship around teaching (Hutchings 1996).

**Enables teaching to be a community endeavor**.  Relatedly, too often in higher education teaching is subject to what Pat Hutchings has called, “pedagogical isolation,” but peer review provides opportunities for us to open our teaching up to a community of colleagues who can nurture improvement (Pat Hutchings 1996).

**Peer review allows for less exclusive reliance on student evaluations.** Student evaluations have become institutionalized in higher education and for the most part provide extremely useful information for the purposes of evaluating faculty, courses, and even entire curricula.  However, students may not always be the best evaluators since they often have limited disciplinary training, they can have biases against certain faculty unrelated to teaching effectiveness, and they can be less cognizant of institutional goals or values than faculty.

**Greater faculty experimentation and rigor.** Just as importantly, an over-reliance on student evaluations in processes of professional review can cause faculty to become overly concerned about receiving positive student evaluations.  In the worst of moments, this can lead faculty to adopt a consumer model of education, shaping our teaching to meet the needs of students over the needs of our disciplines or institutions (Hutchings 1996).  This, in turn, results in faculty becoming overly cautious by refusing to challenge student expectations by using conventional teaching methods, by becoming less rigorous in their standards, and at worst, by feeling a need to entertain more than educate.  Peer review, when done in formative and summative forms alongside student evaluations, can ensure both faculty and students will have a voice in their evaluation, and that faculty have greater autonomy to innovate and to teach rigorously.  This can give faculty the opportunity to focus more intentionally on what helps students learn best, and therefore more directly focus on the quality of their teaching.

**Allows for both formative and summative evaluation.** When done well, peer review involves both formative and summative evaluations.  The inclusion of greater formative evaluation allows for more significant faculty and instructional development by encouraging more critical reflection on teaching and by providing a safer, less risky, and more collegial setting for assessment.

**Improves faculty approaches to teaching.** Daniel Bernstein, Jessica Jonson, and Karen Smith (2000), in their examination of peer review processes found they positively impact faculty attitudes and approaches toward teaching.  While their study did not reveal a necessary shift in faculty attitudes towards student learning and grading, it did change several important aspects of teaching practice.  First, it dramatically impacted in-class practices, particularly the incorporation of more active and collaborative learning, and less reliance on lecturing.  Second, it improved faculty willingness to ask students to demonstrate higher order intellectual and critical thinking skills.  Third, for some faculty it increased the quality of feedback they gave to their students on assignments, and thus improved student understanding and performance.  And lastly, they enjoyed discussing substantive disciplinary and teaching issues with their colleagues, enhancing the scholarly community in their departments and programs.  Peer review therefore shows an ability to improve faculty joy in teaching by improving the relations among faculty and students, and among faculty themselves.

**How to Select Peer Reviewers**

Peer review may take many forms, but usually begins with the selection of peer reviewers drawn most often from within the same department or program as the instructor being reviewed.  The reviewers typically are senior faculty, but sometimes junior faculty as well, who have significant expertise in teaching.  These faculty may be chosen to undertake all peer teaching reviews for the department or program during a specific period, or they may be selected specifically because they share some expertise with the instructor being reviewed.  The person under review also may be granted some choice as to whom one or more of the reviewers may be.  The number of the reviewers may vary but usually include at least two and rarely more than four.

In selecting reviewers, one must be mindful of several criteria**.**

**Institutional Experience.** It helps if reviewers are highly familiar with the department or program, school, and institutional goals, and particularly the processes of peer review itself and the criteria that form the basis of the assessment.

**Integrity.** Peer reviews also function best when reviewers have commitments to integrity, fair-mindedness, privacy, and understanding the reasoning behind the teaching choices of the person under review.

**Trust.** Peer reviewers, especially in formative reviews, work collaboratively with the faculty under review to establish a clear process of evaluation and reporting, therefore peer reviewers who can establish trust are particularly effective.

**Mentorship.** Those under review are particularly vulnerable and often anxious, therefore reviewers who have grace and tact in the process of assessment, can offer feedback with integrity and support, and who can help advise on strategies for faculty development will be most helpful.

**Thorough and Practical.** Peer reviewers should be able to provide summary reports that clearly and thoroughly represent all phases of the process, and that make recommendations that are specific and practical (Center for Teaching Effectiveness, University of Texas, Austin).

**How to Evaluate?**

The peer evaluation itself usually focuses on several aspects of teaching through a process that usually has a series of activities.  The following list of peer evaluation activities represents a sequential, reasonably thorough, and maximal model for peer review, but not all are necessary.

**Develop Departmental Standards for Teaching.** Without a clear set of learning goals for all departmental programs it is difficult to assess teaching with any validity or reliability, and it can leave departments open to biases, inconsistencies, and miscommunications in peer evaluation processes.  One of the greatest benefits of peer reviews of teaching is that it provides an occasion for departments and programs, if not entire schools and universities, to be more intentional, specific, and clear about quality teaching and learning, and the various means to achieve it.  This may be the work of an entire department or a special teaching committee that researches disciplinary and institutional benchmarks and proposes guidelines for review.

**Preliminary Interview.** Peer review processes usually begin with a conversation, sometimes framed as an interview, between the peer reviewers and the teacher being reviewed.  The prime purpose of this is to provide the teacher in question an understanding of the process of peer review, and to offer them the opportunity to provide their input on the process.  The conversation also allows the peer reviewers to begin collecting information about the teaching context, particularly the courses, of the teacher being reviewed.  This context helps to provide better understandings of the teacher’s goals and teaching choices, and may be divided into several dimensions related to the design of their courses (Fink 2005).

**Logistical contexts.** How many students?  Is the course(s) lower division, upper division, a graduate class, etcetera?  How frequent and long are the class meetings?  Is it a distance-learning course?  What are the physical elements of the learning environment?

**Goals.** How have the learning goals of the course(s) been shaped by the department, college, university, or discipline?  Are the courses required or electives?  What kinds of intellectual and skill outcomes is the focus of the course(s)?

**Characteristics of the learners.** What are their ages and other demographic factors that may bear upon teaching?  What is their prior experience in the subject?  What are their interests and goals?  What are their life situations?

**Characteristics of the teacher.** What expertise does he or she have in the subject areas?  What are his or her own assessments of his/her strengths and weaknesses?  What models of teaching did he or she encounter as a student?  What theoretical or practical orientations ground his or her approach to teaching and learning?  What from the teaching and learning scholarship has been influential on his/her teaching?  How do these influences take shape in the teaching of the instructor’s different courses?

**Class Observations.** The goal of the class observations is to collect a sample of information about the in-class practices of teaching and learning.  They typically include two to four class visits to gain reliable data.  If the teacher being reviewed teaches multiple courses, as they often do, the process may involve fewer observations per course (e.g., two).

**What to observe?** The goal is to create a thorough inventory of instructor and student practices that define the teaching and learning environment.  These may vary widely across discipline and teachers, and can be drawn from a broad array of pedagogies, depending on learning goals.  This said, there are several categories of instructor and student practices to note during the observation(s).

**Content knowledge**  
**Use of instructional materials**  
**Class organization**  
**Presentation form and substance**  
**Teacher-Student interactions**  
**Student participation**  
**Assessment practices**

**How to assess teaching practices?** In many institutions, inventories of teaching practices are combined with assumptions about what is conducive to student learning.  It is important for the peer reviewers and the administrators who guide them to be conscious of what they regard as effective teaching and the appropriate evidence for it before committing to an observation process, lest the peer review gather invalid or unreliable data, and lest the process invite peer biases and unexamined pedagogy into the evaluation.  A reasonably representative list of teaching practices, along with more or less explicit value for learning, would include the following:

**Content knowledge**

– Selection of class content worth knowing and appropriate to the course  
– Provided appropriate context and background  
– Mastery of class content  
– Citation of relevant scholarship  
– Presented divergent viewpoints

**Clear and effective class organization**

– Clear statement of learning goals  
– Relationship of lesson to course goals, and past and future lessons  
– Logical sequence  
– Appropriate pace for student understanding  
– Summary

**Varied methods for engagement, which may include…**

– In-class writing  
– Analysis of quotes, video, artifacts  
– Group discussions  
– Student-led discussions  
– Debates  
– Case studies  
– Concept maps  
– Book clubs  
– Role plays  
– Poster sessions  
– Think aloud problem solving  
– Jigsaws  
– Field trips  
– Learning logs, journals  
– Critical incident questionnaire ([see Brookfield](https://static1.squarespace.com/static/5738a0ccd51cd47f81977fe8/t/5750e567f699bbceac6e97f5/1464919400130/CIQ.pdf))

**Presentation**

– Project voice  
– Varied intonation  
– Clarity of explanation  
– Eye contact  
– Listened effectively  
– Defined difficult terms, concepts, principles  
– Use of examples  
– Varied explanations for difficult material  
– Used humor appropriately

**Teacher-Student Interactions**

– Effective questioning  
– Warm and welcoming rapport  
– Use of student names  
– Encouraging of questions  
– Encouraging of discussion  
– Engaged student attention  
– Answered students effectively  
– Responsive to student communications  
– Pacing appropriate for student level, activity  
– Restating questions, comments  
– Suggestion of further questions, resources  
– Concern for individual student needs  
– Emotional awareness of student interests, needs

**Appropriateness of instructional materials**

– Content that matches course goals  
– Content that is rigorous, challenging  
– Content that is appropriate to student experience, knowledge  
– Adequate preparation required  
– Handouts and other materials are thorough and facilitated learning  
– Audio/visual materials effective  
– Written assignments

**Student engagement**

– Student interest  
– Enthusiasm  
– Participation  
– Student-to-student interaction

**Support of departmental/program/school instructional efforts**

– Appropriate content  
– Appropriate pedagogy  
– Appropriate practice

**In-class, formative assessment practices**

– Background knowledge probes, muddiest point exercises, defining features matrix and other “classroom assessment techniques” described in greater detail [here](http://cft.vanderbilt.edu/guides-sub-pages/cats/)  
– Ungraded in-class writing exercises, such as minute papers  
– Discussions  
– Questioning

**Out-of-class, summative assessment practices**

– Class participation  
– In-class writing exercises, graded  
– Presentations  
– Examinations  
– Projects

**Use of observation forms.** To make the process more transparent, reliable, and valid, many departments and programs use observation forms, constructed from items like those listed above, to help peer evaluators track and evaluate teaching and learning practices.  These may include nothing more than checklists of activities; they may provide rating scales (e.g., Likert scales) to assist the evaluation; they may have open-ended prompts that provide space for general commentary and analysis; or, they may involve some combination of all three.  The most thorough forms guide the observer in what exactly they should observe, and prompt them to provide some synthesis and evaluation of their observations.  Several example forms may be found [here](http://www.wayne.k12.ms.us/userfiles/384/PeertoPeerClassroomObservationForm.pdf) (Wayne State University), [here](http://www.albany.edu/teachingandlearning/tlr/peer_obs/New%20Peer%20Observation%20Form.pdf) (New York University, Albany), [here](http://www.academic.umn.edu/provost/peer_review/documents/observation_protocol.docx) (University of Minnesota), and [here](https://cetl.uic.edu/sample-peer-observation-forms/) (University of Illinois, Chicago).

**Evidence of Student Learning.**

**End-of-course student work.** To more thoroughly assess the effectiveness of instruction, peer reviewers may collect evidence of student learning in the form of examinations, written assignments, and other projects from the course of the teacher under review.  Collecting this evidence may be helpful in assessing core competencies expected from the course.

**Student work throughout the course.** Evidence of student learning may be more thoroughly assessed by collecting examples of student work at various times during a course so as to gain perspective on student growth and development.  To do this requires some preparation and lead-time to ensure the teacher under review is sure to collect work from students, and gain their consent for sharing it.

**Grades.** Student grades also may be used as an indicator of student performance, if they are accompanied by contextual information such as a grade distribution, the criteria used to assign those grades, and samples of student work at A, B, C, D, and failing levels.

**Student Evaluations.** In addition to reviewing standard end-of-course evaluations, peer reviewers may choose to solicit letters of evaluation from a sample of students, current or alumni, who have had at least one course with the teacher in question, preferably two or more.  Requesting these from graduates who have a more mature perspective on the effectiveness and impact of the teacher under review can be especially useful.  The request for evaluation letters can be more or less specific in its prompts, but at a minimum typically introduce the importance of the evaluation process for the individual and the institution, and ask for them to assess how effective the teacher was as an instructor, what limitations he or she may have, and what impacts he or she made to their educations.

**Engagement with Centers for Teaching.**If the person under review has attended consultations, workshops, or other programs offered by a campus center for teaching and learning, the evaluation process may consider this to be part of the analysis.

**Advising Activity.** Peer evaluators may wish to make note of the advising activities and load of the teacher in question, along with any special service to the teaching mission of the department, school, or institution.  This may involve some data collection from students the teacher has advised and peers with whom the teacher has collaborated in their teaching service.  For some faculty, this kind of teaching outside typical course structures can be a substantial contribution to the teaching mission of the department.

**Professional Publications, Presentations, and Recognitions.** Peer reviewers also may wish to collect evidence of the scholarly activities in teaching and learning by the teacher in question, such as professional publications, presentations, or awards for their teaching.

**Collaborative Analysis.** Together, each of the activities above provides information that can be assembled into an overall picture of the teacher under review.  After meetings between the peer evaluators to review the data collected, any missing information can be sought and unresolved questions can be answered.  It is then incumbent upon the evaluators to discuss the form and substance of a final assessment and to divide the work of writing it.

**Overall Recommendation.** Typically the written evaluation includes some clarification of the process, the methods, the data collected, and of course any positive feedback and constructive criticism that is necessary, along with suggested improvements.  This will be the substance of a formative or summative assessment by the peer evaluators, one that may be shared with the relevant administrators and the teacher under review, depending on the process adopted.  If the evaluation is formative, this may accompany a series of suggested improvements for teaching and a plan for instructional or curricular development that could include ongoing mentorship, the use of professional development resources such as the Center for Teaching, and further peer evaluation.  If it is a summative evaluation, the recommendation will be used by departmental and university committees and administrators as the basis for a reappointment, promotion, or tenure decision.

**Possible Limitations of Peer Review?**

**Limitations of Peer Observations.** While peer review may be a process that allows for a more rigorous evaluation of a teaching portfolio, it is worth noting that peer observations alone are often insufficient data on which to base an entire teacher’s assessment.  Peer observations represent merely a snapshot of teaching, and thus must be only one component of a teaching portfolio that is subject to peer evaluation, including student evaluations, evidence of student learning, course materials, and self evaluations, just to name a few.

**Bias.** Surely, all methods of teaching evaluation risk biases of one form or another.  One common criticism of peer review processes is that they may invite some bias if they involve limited or unprofessional approaches to information collection and analysis.  This may occur because of several reasons.  Personal relationships between reviewers and those being reviewed can create either hyper- or hypo-critical approaches to evaluation.  Standards of excellence or their application can be highly subjective and individual teaching styles may vary widely, therefore evaluations can be contentious if standards are not defined in advance through rigorous research and open, collaborative processes.  Power relations in departments or programs also can unduly influence open and thorough evaluation.  Other factors may cause peer evaluator bias as well.  Therefore, to avoid the worst cases of bias, peer review must be established via processes that guarantee the greatest rigor, openness, and transparency.

**Collegiality Issues.** Under the best of circumstances, peer review can shape a dialogue about teaching that fosters a teaching community among educators and can lead to more growth-oriented forms of professional development.  However, when it is implemented in less collaborative and more adversarial forms, or when it involves unavoidable consequences such as promotion or job security, anxieties and frustrations can be triggered for both reviewers and those being reviewed.  Therefore peer review must adhere to the highest standards of transparency, integrity, and care for the sake of those under review.

**Time and Effort.** Possibly the most common critique of peer review processes, and the reason they are not more commonly used in the academy, is that they require significant time and effort.  Departmental and campus administrators must define the process, establish standards, train and prepare reviewers, perform peer observations, review portfolios, draft assessments, and have multiple dialogues with those under review.  Each step requires preparation if it is to be fair, transparent, and professional.  Any shortcut may compromise the rigor, care, or goals of the evaluation.  However, there are several shortcuts each with potential costs.

**Rely on the expertise of senior colleagues, administrators, and the Center for Teaching.** There are typically those on campus that my have sufficient knowledge to assist in defining departmental learning or teaching goals, in determining what data to include in a teaching portfolio, in training peer observers, in drafting assessments, etcetera.  These sources of expertise may be helpful in streamlining the process with little cost to its integrity, as long as their suggestions may be tailored to the needs of the department or program in question.

**Use predefined standards for teaching and learning.** Rather than spend significant time adjudicating which learning and teaching goals are appropriate, department or program leaders may decide to use existing language in university or departmental missions, course catalogs, accreditation reports, other constituting documents, or the operating principles of the Center for Teaching.  This may grant some efficiency with limited costs to the integrity of the peer review process.  However, vague and imprecise learning goals that sometimes characterize constitutional documents (e.g., “critical thinking”) may be of little help in benchmarking a specific set of courses or teaching strategies.  Likewise, departments and programs may have particular teaching challenges that broad standards may not take into consideration.  Both difficulties can leave departments or programs open to unclear standards, unfair or inconsistent judgments, and miscommunications.

**Collect data judiciously.** One of the more time consuming tasks of peer review is combing through all facets of a teaching portfolio, particularly if it includes samples of student work.  To save time, some peer review processes rely largely upon peer observation, in addition to student evaluations of teaching, and do not collect teaching portfolios or examples of student work.  Others collect only limited samples of student work, such as grade distributions and examples of student work at A, B, C and D levels to evaluate an instructor’s assessment and grading strategies.  Other data collection short cuts may be possible as well.  However, more limited data may allow fewer contextual interpretations of a teaching career, and peer observations alone are merely in-class snapshots of instructional performance, not a more encompassing perspective on all phases of teaching.  These may lead a department or program to make less informed and fair judgments.

**Use templates for written peer evaluation reports.** Final written reports need not be highly expansive analyses, but may represent more of a thorough check list with brief sections of commentary on challenges and successes that become points of discussion between peer reviewers and the instructor under review.  This form or report can save valuable time, but it also may provide limited feedback to the instructor under review, possibly affording him or her less useful guidance on where to improve his or her teaching.

**Only summative evaluation.** A department or program may limit peer evaluation to only summative and not formative assessments of teaching.  This would limit opportunities for faculty development, hinder data collection, create more tensions between reviewers and those being evaluated, and thwart the formation of collegial cultures that improve teaching for entire departments and programs. However, many departments and programs have used this shortcut to conduct peer review.

**Concluding Thoughts**

Peer review of teaching, when done well, has many benefits in fostering teaching excellence, creating collegial communities of scholar teachers, and more fair and transparent cultures of professional development.  By contrast the challenges of peer review, while not insignificant, are small by comparison.  Peer review of teaching, as in research, enhances the integrity and innovation of teaching and is a practice whose institutionalization is long overdue.

**Bibliography**

* Bernstein, Daniel J. 2008. “Peer Review and Evaluation of the Intellectual Work of Teaching.” Change. March/April.
* Bernstein, Daniel J., Jessica Jonson, and Karen Smith. 2000. “An Examination of the Implementation of Peer Review of Teaching.” New Directions for Teaching and Learning. 83: 73-86
* Bernstein, Daniel., A.N. Burnett, A. Goodburn and P Savory. 2006. *Making Teaching and Learning Visible: Course Portfolios and the Peer Review of Teaching*. Anker.
* Center for Teaching Effectiveness. “Preparing for Peer Observation: A Guidebook.” University of Texas, Austin.
* Chism, Nancy V. 2007. *Peer Review of Teaching: A Sourcebook*. *2nd Edition.* Anker.
* Glassick, C. M. T. Huber, and G. Maeroff. 1997. *Scholarship Assessed: Evaluation of the Professoriate*. Jossey-Bass.
* Hutchings, Pat. 1995. *From Idea to Prototype: The Peer Review of Teaching*. Stylus
* Hutchings, Pat. 1996. “The Peer Collaboration and Review of Teaching.” ACLS Occasional Paper No 33.
* Hutchings, Pat. 1996. *Making Teaching Community Property: A Menu for Peer Collaboration and Peer Review*. Stylus
* Hutchings, Pat. 1998. *The Course Portfolio*. Stylus
* Perlman, Baron and Lee I. McCann. 1998. “Peer Review of Teaching: An Overview.” Office of Teaching Resources in Psychology and Department of Psychology, Georgia Southern University.
* Seldin, P. 1997. *The Teaching Portfolio*. *2nd Edition*. Anker.
* Seldin, P. 1999. *Changing Practices in Evaluating Teaching: A Practical Guide to Improved Faculty Performance and Promotion/Tenure Decisions.* Jossey-Bass.
* Shulman, Lee S. 2004. *Teaching as Community Property: Essays on Higher Education*. Jossey-Bass.

[Creative Commons License](http://creativecommons.org/licenses/by-nc/4.0/)  
This teaching guide is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](http://creativecommons.org/licenses/by-nc/4.0/).