Report of the Joint Faculty Council-Chancellor Task Force on the Evaluation of Teaching at Rutgers University-New Brunswick

discussion draft

March 29, 2017
table of contents:

Chapter 1: Why does the evaluation of teaching matter?

Chapter 2: Review of evaluation of teaching
         at Rutgers University-New Brunswick and other universities

Chapter 3: Suggested frameworks for evaluation of teaching
         at Rutgers University-New Brunswick

Chapter 4: The role of student surveys in the evaluation of teaching

Chapter 5: What’s next?
Chapter 1: Why does the evaluation of teaching matter?

At Rutgers University, a prominent research institution, teaching matters. One of the key provisions in the University’s mission statement says that Rutgers will provide “for the instructional needs of New Jersey’s citizen through its undergraduate, graduate, and continuing education programs.” Students who attend Rutgers rightly expect a high-quality education, with excellent teachers. As an AAU member institution, at Rutgers, knowledge operationalized through research and scholarship is of crucial importance. The dissemination of knowledge through teaching is an inseparable and essential component of scholarship, and is thus a primary mission of the university. For tenure track (“TT”) faculty, a mediocre teaching record can jeopardize or weaken the case of one’s chance of promotion, even for someone with a strong publication record. For non-tenure-track (“NTT”) faculty – who teach roughly two-thirds of all undergraduate credit-hours – teaching is especially important as it is their primary task at the University. For part-time lecturer (“PTL”) faculty, who teach on a semester-to-semester basis with no guarantee of renewal, a mediocre teaching record can jeopardize one's chances of continuing at Rutgers.

In sum, because teaching – good teaching – matters, it follows that the evaluation of teaching is crucial. It is very important that teaching be fully, fairly, and accurately assessed.

The evaluation of teaching at Rutgers has been extensively discussed for many years, and several plans for evaluation of teaching have been put forward in the past. However, although teaching matters very much at Rutgers, there is a longstanding consensus that the evaluation of teaching is not as well-organized, systematic, thorough, or appropriate as it should be. This recognition led to discussions during the spring and summer of 2016 between the New Brunswick Faculty Council and the New Brunswick Chancellor, Dr. Richard Edwards, and to the formation of the task force whose report and recommendations for the evaluation of teaching at the Rutgers University—New Brunswick (“RU-NB”) are set out in this document.

The members of the task force (listed in Appendix A) are drawn from a variety of disciplines and roles in RU-NB. Some of us are NTT faculty; others are TT faculty; still others are administrators; one of us is an undergraduate student. Each of us has brought a particular perspective on the evaluation of teaching. However, each of us considered that we were serving on this task force not as representatives of any particular unit within the University, but, rather, as persons whose expertise, experience and judgment would help in developing new procedures to improve the evaluation of teaching throughout RU-NB.

Although we believe that this report will contribute to better evaluation of teaching within RU-NB as a whole, we are well aware that it would be unwise and unrealistic to attempt to apply uniformly a single set of procedures for evaluating teaching throughout RU-NB: RU-NB is simply too diverse and complicated for that to be either reasonable or appropriate. Because RU-NB is a highly diverse institution – and, rightly, takes pride in its diversity – we know that there is no way that one single set of evaluation procedures could sensibly and appropriately be used to evaluate teaching by all faculty, in every field of study, in classes of all different sizes and types. In what follows, we take what we believe to be the most appropriate approach: to provide for a certain amount of basic uniformity for the evaluation of teaching at
all of the units within RU-NB, while still allowing considerable freedom and flexibility for each
unit to take into account its particular circumstances and, thus, to make adjustments to best
accommodate its own needs.

Finally, it is important to note that evaluation of teaching serves a variety of purposes and
needs for different groups within the university. The availability of reliable information about
individual professors’ teaching is important to senior administrators who must make decisions
about promotion, tenure, and teaching awards. Individual faculty members, and their
departments’ graduate and undergraduate directors, need reliable evaluations of teaching for
diagnostic and remedial purposes. Evaluations of teaching provide students with an opportunity
to express their opinions about instruction, and give students an opportunity to consult such
evaluations as they select courses. In making suggestions for evaluation of teaching, we have
been mindful of the needs of each of these different groups within RU-NB.

The plan of this report is as follows. In the next chapter, we provide an illustrative
(though not exhaustive) survey of procedures for evaluating teaching, first at a variety of
different units within RU-NB and then at number of other universities, including most of the
universities in the Big Ten. In Chapter 3, we provide a general outline of procedures for
evaluating teaching for consideration by units in RU-NB; since the circumstances of different
types of faculty (e.g., TT and NTT faculty) are quite different, our discussion here individually
considers treats each category of faculty separately. In Chapter 4, we discuss student course
surveys. We conclude in Chapter 5 by discussing our recommendations for improving the
evaluation of teaching at RU-NB, and by describing procedures for implementing these
recommendations.
Appendix A: Membership of the Joint Faculty Council-Chancellor Task Force
on Evaluation of Teaching at Rutgers-New Brunswick

Tisha Bender: Assistant Teaching Professor, English Writing Program, SAS
David Chapman: PTL, Music, Mason Gross School of the Arts
Francine Conway: Dean, Graduate School of Applied and Professional Psychology
Linnea Dickson: Assistant Teaching Professor, Psychology, SAS
Amy Finkelstein: Senior, Rutgers Business School
Gary Gigliotti: Associate VP, Academic Affairs (Center for Teaching Advancement and Assessment Research)
Drew Gitomer: Professor, Graduate School of Education
Roseli Golfetti: Assistant Teaching Professor, Office of Undergraduate Instruction, Division of Life Sciences, SAS
Robert Heffernan: Vice President, Office of Institutional Research and Academic Planning
Michael LaSala: Associate Professor, School of Social Work
Steve Peterson: PTL, Mathematics, SAS
Camilla Stevens: Associate Professor, Spanish and Portuguese, SAS
Mark Killingsworth: task force chair; Professor, Economics, SAS
Chapter 2: Review of evaluation of teaching at Rutgers University-New Brunswick and other universities

In this chapter, we discuss the ways in which teaching is evaluated at Rutgers University-New Brunswick and elsewhere. It will quickly become apparent that, both here and elsewhere, teaching is evaluated through a great variety of means, and that few universities or departments are quite like any other in terms of the techniques they use to evaluate teaching.

A. Instructional faculty at Rutgers University-New Brunswick: An overview

Before we discuss the various ways in which teaching is evaluated at Rutgers University-New Brunswick (“RU-NB”), it is important to set the stage by noting that there are several distinct categories of instructional faculty at RU-NB: tenured and tenure-track faculty (“TT”); full-time non-tenure-track faculty (“NTT”); and part-time lecturers (“PTL”); and teaching assistants. The terms and conditions of employment for these different categories of faculty are quite different. These differences have an important bearing on how – and even whether – these different kinds of faculty are evaluated in terms of their teaching. For further details on these different kinds of faculty, please see Appendix A.

B. Evaluation of teaching at Rutgers University-New Brunswick

At RU-NB, schools (e.g., the School of Arts and Sciences, the School of Business) and their constituent departments determine how teaching is evaluated. Most schools at RU-NB use end-of-course surveys, typically (though not always) administered through the Student Instructional Rating System (SIRS). These end-of-course surveys use Likert ratings (typically based on, e.g., a four- or five-point scale) and free-form student comments. Some departments engage in peer review and direct classroom observation for tenure-track faculty seeking tenured appointments; less frequently, departments also use such methods in evaluation tenured faculty subject to post-tenure review. However, we are not aware of any University-wide policies requiring the evaluation of teaching for PTLs or NTTs. We are also not aware of requirements for evaluation of teaching in connection with post-tenure review of tenured faculty.

Schools and departments use a variety of other processes and methods to evaluate their faculty members’ teaching. Pedagogical approaches to teaching evaluation include the scholarly examination of teaching methods; assessments of student learning (for example, assessments of projects or support for undergraduate and graduate research); and the development and review of teaching portfolios (self-reflective teaching statements, statements about contributions to instruction, samples of syllabi, exams and quizzes, student work, and other course-related materials; and videos of teaching). Direct methods of teaching evaluation may include peer review of instructional materials, administrators’ ratings of teaching and review of student work samples. Indirect evaluations of teaching may include teaching awards and mentoring of new faculty. School and department-wide assessments for program evaluation may include focus-group interviews of students, exit interviews of graduating students, alumni surveys and employer ratings of student graduates.

1 The evaluation of teaching at Rutgers University has been discussed over the years; see, e.g., Appendix B.
The diversity of disciplines, instructional methods, class sizes, and learning goals requires flexibility and autonomy in the evaluation of teaching. However, approaches to evaluating teaching that serve as models of best practices at RU-NB are exemplified by the ways in which a number of departments, including the departments of History and English, go about this task.

1. **Evaluation of teaching in the Department of History**

In the History Department, new faculty are assigned a mentor who supports their transition to their role and to the teaching environment. The mentor also assists the new faculty in understanding the department’s promotion criteria and process.

Prior to the faculty’s reappointment/tenure there are at least two opportunities for classroom observations conducted by two faculty members. Peers also review syllabi and assignments, and discuss the class performance orally and in writing with the faculty member.

During promotion meetings, the faculty teaching reports are reviewed and inform tenure decision discussions. The teaching reports integrate classroom visit reports, the SIRS materials, and any extra evaluations that professors might have used (usually a paper evaluation) that are designed to assess student learning in history.

Tenure reviews and mentorship of faculty teaching continue after tenure, but occur less regularly than is usual before the tenure decision.

2. **Evaluation of teaching in expository writing (English 101)**

Expository Writing (English 101) is the required writing course for all students at Rutgers University in New Brunswick, and it is usually completed in the first semester. As the course description notes, the course is “designed to prepare students for the writing they will do at the university and in their professional lives. Both at school and at work, writing usually involves three fundamental activities: reading articles, reports, or books intended for an educated audience; making connections among multiple sources; using this knowledge to develop an independent thesis that responds to the ideas of others.”

In this course, students read a complex text – usually a non-fiction text in the context of current issues – and then write five papers that build consecutively from the development of independent claims to an argument. The goal of the course is to teach students to be active participants in society, broaden the scope of their thinking, and develop and progress to independence in writing and revising their thinking about issues. Simply put, the course trains students to be active participants in their own learning.

The assessment of learning is guided by the course’s structure which includes training, mid-semester reviews, mentor program, support for teaching assistants, oversight by course directors and coordinators and course evaluations. It is important to note the evaluation of teaching is a natural and continuing outgrowth of this structure.

---

2 See [http://wp.rutgers.edu/component/content/article/60-course-listing/55-355101](http://wp.rutgers.edu/component/content/article/60-course-listing/55-355101).
Specifically, the department provides a detailed almanac\(^3\) that provides training for course instructors. The training includes peer review exercises, exercises in developing course syllabi, and preparation of instructional materials and lectures for the first 4 weeks of class. This instructor training and mentoring is particularly important for the professional development of many of the part-time lecturers who make up a large part of the instructional staff in this program.

Each instructor participates in a Mid-Semester Folder Review, meeting with the department director for an hour or more. Together they review the work of the instructor’s class. The review process offers the instructor support in some key instructional areas, including assistance in regulating their grading (a grading rubric is provided) and in managing potential concerns about students – for example, how to assist under-performing students, address plagiarism, and help students to meet standards.

The End-of-Semester Folder Review is a final review of the instructor’s teaching. The Expository Writing course has a large number of instructors with various backgrounds and teaching experiences. The folder review sessions are opportunities for the director to assess the teaching quality of each instructor and to assess the standards for grading in such a huge number of course sections and students. The director of the folder review has the final say concerning the grades assigned in each section, providing for standardizing or normalizing grades across the multiple sections.

The program provides part time lecturers (PTLs) with mentors. The goal of the mentor program is to help the PTLs to develop a relationship with the department, facilitate their connection to the department, and provide them with an advocate. Lecturers and mentors meet in person and/or have phone contact three times per semester. The lecturer’s teaching is also observed by their mentors. The teaching observation is designed not as a means of evaluating the instructor’s performance, but rather as a way to provide the instructor with helpful feedback.

The department has a tradition of offering a six-session required teaching seminar for teaching assistants. An additional seventh session is optional.

Most courses have a course coordinator (especially those courses with multiple sections). The course coordinators are charged with developing curriculum and teaching manuals for the courses; and disseminating the course material to the instructors. The coordinator also mentors instructors and facilitates orientation for new and returning instructors. Orientation is generally held at the beginning of each semester.

Overall, the structure of the program provides an ongoing evaluation of teaching. The department views training and support for instructors as integral to the evaluation of teaching. For example, although the ‘Folder Review Process’ is evaluative, the key to a successful review is the training and support for the instructor provided by the department. The administrators who act as mentors and coordinators highly value their interactions with the instructional staff and find the folder reviews a key element in improving the teaching of individuals and the program at large.

---

Students evaluate the course through the widely-used Student Instructional Rating Survey ("SIRS"). Students also use informal online review sites to rate their experiences in the course and provide support to future students. The department’s view is that although the course has a reputation for being difficult, generally students rate the courses positively and course instructors find the teaching experience gratifying.

3. **Evaluation of teaching by non-tenure track faculty in the School of Engineering**

In its “School of Engineering Criteria for Evaluation, Reappointment and Promotion of NTT Faculty,” the School of Engineering sets out its expectations for evaluation of teaching performance on the part of non-tenure track faculty:

Teaching performance may be evidenced by documents, videos and/or other summaries appropriate to their assigned responsibilities. Evidence might include a teaching portfolio, student ratings, curriculum development, laboratory manuals, student work products, assessment performance indicators and assessment rubrics, published archival papers, textbooks, online modules, conference presentations, student research supervision, and evidence of achievement of student learning objectives, for example. Evaluation of teaching performance should also include evidence of a faculty member’s knowledge and use of current advances in pedagogy, breadth and depth of knowledge in the subject area appropriate to their NTT rank, and effective communication skills.4

4. **Evaluation of teaching by teaching assistants**

The SAS Anthropology Department uses a brief series of questions for faculty to assess their teaching assistants (TAs), addressing their duties, observation, areas of strength, overall rating and suggestions for improvement.5

The SAS Mathematics Department uses an online “Recitation Instructor Evaluation Form” that asks questions about the technical aspects of the instructor’s teaching performance; the mathematical content of the instructor’s teaching performance; the instructor’s engagement of students; assistance in grading and constructing exams and quizzes; other comments and recommendations.6 The Mathematics Department also provides a resource page containing undergraduate instructional information. Other resource pages are only available to those in the Mathematics Department with an authentication code.7

5. **Evaluation of teaching in the Rutgers New Jersey Agricultural Experiment Station (NJAES) Cooperative Extension**

Cooperative Extension, an informal science and experiential outreach to the general community, uses research-based knowledge in developing its programs. Faculty and professional

---

4 See [http://soe.rutgers.edu/sites/default/files/imce/Reappointment%202015.pdf](http://soe.rutgers.edu/sites/default/files/imce/Reappointment%202015.pdf)
5 See [http://anthro.rutgers.edu/downloads/graduate/graduate-forms/12-facultyevaluationoftas/file](http://anthro.rutgers.edu/downloads/graduate/graduate-forms/12-facultyevaluationoftas/file).
7 See [https://www.math.rutgers.edu/undergrad/Dept/facinfo.html](https://www.math.rutgers.edu/undergrad/Dept/facinfo.html).
staff assess their teaching in the short term with end-of-presentation surveys, a customized version of the SIRS, as well as the medium and long term impacts on their audiences.  

Faculty who teach in Cooperative Extension programs are provided with training and resource materials for skills development. There is an extensive collection of program evaluation resources with evaluation forms for assessing teaching effectiveness with respect to the many different kinds of “students” in their seminars, presentations and programs. Their goal for teaching assessment and program evaluation is to develop the instructors’ abilities and skill to improve the evaluation of how they teach a wide variety of participants in their programs, i.e., businesses, government agencies, youth and adult residents in New Jersey. Their materials address all aspects of assessment including pedagogy, theory and practice. County Agents are primarily assessed on their extension practice for reappointment, tenure, and promotion by their supervisors and administrators of the Cooperative Extension programs.

3. Evaluation of teaching in the Department of Physics and the School of Arts and Sciences Office of STEM Education

In evaluating instruction, the Department of Physics uses a variety of methods. Peer evaluation is carried out by one or two faculty members of the departmental peer teaching evaluation committee, and is coordinated by the chair of the committee. The committee members are chosen based on their teaching skills and experience. The frequency of evaluation is typically the highest for junior faculty and can be as high as thrice before the tenure case. Tenured faculty are typically evaluated whenever promotion is a possibility. The frequency thus decreases with seniority, but student ratings, student feedback and undergraduate director input can play a role in peer evaluations at any stage.

The evaluator reviews the class logistics such as enrollment, textbook, and website, and arranges a class visit with the instructor, giving the instructor leeway in choosing the topic for the class. The visit report considers numerous factors, such as the appropriateness of the instruction level, engagement of and with students, blackboard usage, attendance and so on. Typically, the visit report contains concrete comments and suggestions for the instructor; if there is a second faculty evaluator, that person will be alerted to any deficiencies and suggestions made as a result of the first visit. The visit report also includes a discussion with the instructor of general issues and experience with the course, teaching philosophy etc. The report goes into the instructor's permanent record along with the instructor's responses, if any. The report is shared with the undergraduate and graduate directors in the department. The process is cooperative and fully open to the evaluated faculty at all stages.

Evaluations of NTT faculty are carried out in a similar fashion. These have started only recently and thus far have been on an ad-hoc basis.

In addition to these general procedures for evaluation of teaching, research on STEM education and effective teaching has been a significant part of the work of faculty in Physics with colleagues in Mathematics and Education. For several years, the Physics and Astronomy

---

8 See [http://njaes.rutgers.edu/evaluation/](http://njaes.rutgers.edu/evaluation/).
9 See [http://njaes.rutgers.edu/evaluation/resources/](http://njaes.rutgers.edu/evaluation/resources/).
Education group (PAER) and the Area Dean’s Office in the School of Arts and Sciences have been studying evidence-based teaching methods and methods for the assessment of teaching in STEM fields. The Office of STEM Education was created and with faculty from PAER, Chemistry and Life Sciences departments.10

Together with innovative faculty in Mathematics, Education, Chemistry and Life Sciences, PAER faculty have undertaken course transformation projects wherein a variety of modern research-based teaching strategies are employed to improve student learning. Assessment of teaching in these projects then requires different approaches for evaluation of teaching which include peer observation, elements of teaching portfolios and a focus on authentic learning demonstrated by evidence in student work products.

C. Evaluation of teaching at Big Ten universities and elsewhere

To assess the kinds of materials and metrics used in the evaluation of teaching in peer institutions, we conducted a survey of teaching evaluation at Big Ten universities. A copy of the survey report is attached in Appendix C.

The institutions that responded to the survey were Chicago, Illinois, Indiana, Iowa, Michigan, Michigan State, Minnesota, Northwestern, Ohio State, Penn State, Purdue, and Wisconsin. Indiana University-Purdue University-Indianapolis (“IUPUI”) sent their rubrics for evaluation of teaching; the Interim Associate Vice chancellor for Faculty Affairs at the University of Nebraska sent a brief letter description of the evaluation of teaching process. (See Appendix D.)

The survey results show that these universities use a wide range of practices in evaluating teaching. Not surprisingly, two particular kinds of teaching evaluation are predominant: end-of-course numerical ratings derived from end-of-course surveys of students, and written student comments gathered as part of those surveys: see Table 1.

---

Table 1. Evaluation of Teaching Practices

<table>
<thead>
<tr>
<th>Q2 - Do your Evaluation of Teaching practices include any of the following? (Please check all that apply.)</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>End-of-course ratings surveys (i.e. instructor ratings and course review)</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>Use of written comment from end-of-course ratings surveys</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>Peer review with class observations</td>
<td>83.33%</td>
<td>10</td>
</tr>
<tr>
<td>Self-reflective statement</td>
<td>75.00%</td>
<td>9</td>
</tr>
<tr>
<td>Teaching awards</td>
<td>75.00%</td>
<td>9</td>
</tr>
<tr>
<td>Support of undergraduate student research</td>
<td>75.00%</td>
<td>9</td>
</tr>
<tr>
<td>Mentoring of new faculty in teaching</td>
<td>66.67%</td>
<td>8</td>
</tr>
<tr>
<td>Scholarship of Teaching</td>
<td>66.67%</td>
<td>8</td>
</tr>
<tr>
<td>Support of graduate student research</td>
<td>66.67%</td>
<td>8</td>
</tr>
<tr>
<td>Peer review of instructional materials</td>
<td>58.33%</td>
<td>7</td>
</tr>
<tr>
<td>Mid-course course survey</td>
<td>58.33%</td>
<td>7</td>
</tr>
<tr>
<td>Contributions to instruction</td>
<td>58.33%</td>
<td>7</td>
</tr>
<tr>
<td>Administrator ratings of teaching, scholarship and service</td>
<td>58.33%</td>
<td>7</td>
</tr>
<tr>
<td>Assessments of student learning</td>
<td>50.00%</td>
<td>6</td>
</tr>
<tr>
<td>Teaching Portfolio</td>
<td>41.67%</td>
<td>5</td>
</tr>
<tr>
<td>Focus-group interviews of students</td>
<td>41.67%</td>
<td>5</td>
</tr>
<tr>
<td>Exit interviews of graduating students</td>
<td>33.33%</td>
<td>4</td>
</tr>
<tr>
<td>Review of student work samples</td>
<td>33.33%</td>
<td>4</td>
</tr>
<tr>
<td>Alumni surveys</td>
<td>25.00%</td>
<td>3</td>
</tr>
<tr>
<td>Videos of teaching</td>
<td>25.00%</td>
<td>3</td>
</tr>
<tr>
<td>Employer ratings of student graduates for program evaluation</td>
<td>25.00%</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>12</td>
</tr>
</tbody>
</table>

Our survey results indicate that the techniques and procedures for evaluation of teaching vary substantially not only from university to university, but also from department to department within the same university. Many of the respondents commented, “All of the [evaluation] items are not used by all faculty,” or that “not ALL departments do ALL of these things.” There is a great deal of variability and flexibility afforded to the schools and departments within schools: “methods vary considerably from college to college and even between departments in a given college”, it is “up to the College to determine” and “every Tenure-initiating Unit (department or college) sets its own policy within guidelines.” There are, however, some common items, on which Penn State’s comments are fairly representative: “Our [promotion and tenure] process requires: (1) Peer Review (course observation) (2) Student Feedback (a) Student Ratings of Teaching Effectiveness (Likert items) and (b) one other form of feedback from students.”

In most institutions, consideration of instructional methods was an important element in the evaluation of teaching:
Table 2. Diversity of Instructional Methods

<table>
<thead>
<tr>
<th>Q4 - Does your Evaluation of Teaching process take into account the diversity in instructional methods? (Please check all that apply.)</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Studio courses</td>
<td>100.00%</td>
<td>9</td>
</tr>
<tr>
<td>Lecture based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>Discussion based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>Performance based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>Clinical work</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>Hybrid courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>Online courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>Case study courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>Small group interaction or seminar courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>Laboratory courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>Fieldwork courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>Practicum</td>
<td>55.56%</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>9</td>
</tr>
</tbody>
</table>

Across the Big Ten there is an important awareness of diversity of instructional methods influencing measures of teaching effectiveness. Again the responses highlighted the institutional variability and flexibility for colleges and departments to attend to the variety of instructional styles warranted by the type of course. Comments include:

- “the various colleges and departments use a variety of tools and methods to evaluate teaching processes - there is a great deal of autonomy across the institution.”
- “We recognize the diversity of methods, and -- particularly in evaluating faculty at transition points (promotion) -- departments are encouraged to evaluate them accordingly.” “peer reviewers do not expect the same from faculty teaching 800 students as they do from faculty teaching a studio course,”
- “Methods used for these different instructional methods may vary, based on the type of course and the department.”
- “Almost all of our units use a locally developed student feedback form for most didactic classroom and online experiences.”
- “Of course every unit uses different elements beyond the student survey, although most differ in ways that are not related to instructional methods, but more the culture of the discipline and department.”
- “We do have departments which have created different cores for different types of courses such as labs, discussion.”
- “On our online courses, we have some additional surveys that are sent out to gather more detailed information.”
Most have some mention of performance based on type of course. As one respondent put it, “most departments will note differences in student ratings for different course types (e.g., ‘she did great in seminars and small group sessions but she was perceived as teaching poorly in large lecture classes’).”

Table 3 summarizes other activities that are sometimes considered in an evaluation of teaching of faculty:

<table>
<thead>
<tr>
<th>Q6 - Does your Evaluation of Teaching process consider these additional components? (Please check all that apply.)</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Teaching</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>Development of curricula, new courses and course materials</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>Workload of classroom instruction (class size, course level, mode of instruction)</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>Teaching assistant supervision</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>Advising students in the major</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>Independent study courses</td>
<td>44.44%</td>
<td>4</td>
</tr>
<tr>
<td>Advising undeclared major students</td>
<td>33.33%</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>9</td>
</tr>
</tbody>
</table>

Many of the Big Ten institutions mention additional components that may be used in assessment of teaching effectiveness. Some respondents indicated that this is often done as part of the narrative on teaching and vary with department and faculty role. Comments include:

- “This is highly variable depending on department/unit. Each of our units has a document guiding evaluation processes within that unit.”
- “I am certain that the degree to which they are discussed would vary tremendously across the institution.”
- “Advising is considered part of service. Independent study courses would be listed, but they do not count as part of faculty load. The department head might commend a faculty member for developing curricula, but it is unlikely to "count" much overall.”
- “I don’t know how one would evaluate teaching/learning without considering the full array nature of the course.”
- “Our faculty core dossier asks for a full listing of credit and non-credit teaching, including individual instruction and research supervision, as well as a discussion of course and curriculum development.”
- “Again, it varies by department and school to the extent that these "count." It also depends on the type of instructor (tenure line or teaching line).”
Table 4. Use of Rubrics for any components

<table>
<thead>
<tr>
<th>Q8 - Do you use rubrics for any components of your Evaluation of Teaching process?</th>
<th>Answer</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>27.27%</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>72.73%</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

A few of our colleagues in the Big Ten do use rubrics for peer observation.

- “We have both Student Ratings of Teaching (SRTs) which are standard questions, as well as resources for peer review of teaching. Link: http://www.academic.umn.edu/provost/peer_review/index.html
- “I don’t know, but probably not. Not unless they use a form we have created for peer review, e.g. http://www.schreyerinstitute.psu.edu/pdf/Classroom_Observation_Checklist_Form.doc
- “I would say that we don’t often use them unless we develop them for particular courses (and by request of the instructor). When we observe teaching, however, we might use a rubric.” “Individual departments may use rubrics. I am only aware of some teaching observation/peer observation protocols.”

Table 5. Allow for Reflection on Development

<table>
<thead>
<tr>
<th>Q10 - Do your Evaluation of Teaching practices allow for instructors to reflect on their development over time?</th>
<th>Answer</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>90.91%</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>9.09%</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

In most tenure and promotion processes across the Big Ten, there are opportunities for candidates to add a statement about teaching to a teaching dossier, narrative or portfolio.

- “For every review process (and particularly at promotion and tenure time), instructors are asked to reflect on their teaching.”
- “Not unless they are allowed to include philosophy statement or teaching portfolio as supplemental at the department-review level.”
- “Faculty may make statements in dossiers that show growth over time. We encourage this reflection.”
- “Our instructors are given access to their end of semester scores. They can also request a longitudinal report of their performance on scores to match it with any changes in teaching methods they have used. This, along with support from our Instructor Development team, provides one method to improve teaching.”
- “Tenure line faculty submit letters related to their reflection on their development; teaching line faculty are expected to reflect and improve/develop over time.”
“Instructors are free to pursue formative feedback and free to include whatever they like in their P & T documentation.”

So, as with the other questions in the survey, there are places where faculty can add personal reflections on their teaching and development, but where and how varies across those surveyed and within their institutions.

Table 6. Use of End-of-semester surveys

<table>
<thead>
<tr>
<th>Q12 - Are end-of-course surveys used in your institution or department for any of the following? (Please check all that apply.)</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improving teaching</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>Self-evaluation</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>Reappointment of non-tenure track faculty (faculty of the practice or clinical faculty)</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>Tenure decisions</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>Promotion decisions</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>Promotion of non-tenure track faculty (faculty of practice or clinical faculty)</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>Reappointment of tenure-track faculty</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>Reappointment of casual or adjunct instructors</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>Reappointment of teaching assistants</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>end-of-course surveys are not used for personnel decisions</td>
<td>0.00%</td>
<td>0</td>
</tr>
<tr>
<td>end-of-course surveys are not used at this institution</td>
<td>0.00%</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>12</td>
</tr>
</tbody>
</table>

All of the Big Ten respondents use the end-of-semester student surveys for formative and summative personnel decisions. Our respondents noted:

- “The primary point of these evaluations is to improve teaching. If instructors are not inherently motivated, our use of these surveys for reappointment, tenure, and promotion helps motivate them.”
- “The end-of-term ratings are used for a variety of purposes, both formative and summative. Formatively, individual faculty use the feedback to inform their future teaching. The ratings are also used as one component of summative evaluations at all levels.”
- “Penn State is one of the few institutions at which the ‘Student Ratings of Teaching Effectiveness’ were developed by the Faculty Senate for the purposes of tenure, promotion, and annual review. All contingent faculty (non-tenure track teaching faculty, full or part-time, and pre-tenure faculty) must have the SRTEs administered in every course. Each college has developed policies about the frequency of SRTE administration for tenured faculty. But students have come to expect to evaluate faculty in every course.”
- “Any of the checked items can be used--and in some cases, are required to be used – as part of the portfolio of teaching submitted for tenure, promotion, awards.”
- “The weight given to end-of-course surveys varies widely by department and school.”
“Student feedback, usually limited to end of term surveys, is a major portion of most evaluation of teaching. We urge a broader range of data.”

“The units make the decisions to the extent that they rely on end of course surveys for evaluation.”

“Our end of term ratings have multiple purposes: (a) faculty can use them to improve/learn about what worked in their classes etc.; (b) administrators will use them for promotion and retention decisions and to benchmark across courses, sections, instructors etc.; (3) students use them to tell other students about the class (they are online and visible to everyone at the institution).”

It seems that among the Big Ten institutions there is conformity in that the surveys used at the end of the semester are the responsibility of the University to develop, although there are contributions by subordinate units for specific needs of the college or department.

Table 7. Unit developing surveys

<table>
<thead>
<tr>
<th>Q14 - The end-of-course surveys were developed by.... (please check all that apply)</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>the University</td>
<td>100.00%</td>
<td>11</td>
</tr>
<tr>
<td>the Department</td>
<td>54.55%</td>
<td>6</td>
</tr>
<tr>
<td>the School</td>
<td>54.55%</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>11</td>
</tr>
</tbody>
</table>

Over all the respondents, there is a high level of use for the end-of-semester surveys.

Table 8. Units using surveys

<table>
<thead>
<tr>
<th>Q15 - The end-of-course surveys are used by....(please check all that apply)</th>
<th>Percent</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>the Department</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>the School</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>the University</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>12</td>
</tr>
</tbody>
</table>

In sum, procedures used in the evaluation of teaching across the Big Ten institutions typically include end-of semester student ratings surveys (100%), the use of written comments from those surveys (100%) and the use of peer review based on classroom observation (83.33%). The variety of instructional methods are considered in reviews of teaching effectiveness: for Studio courses 100% of the schools, Lecture, Discussion, Clinical, Hybrid and Online courses in 88.89% of the reporting institutions.
Other contributions to teaching, i.e. team teaching and development of curricula, new courses and course materials are included in teaching evaluation by 88.89% of our Big Ten peers.

Indiana University-Purdue University-Indianapolis sent their rubrics for evaluation of teaching, see Appendix D. In some of the other schools rubrics are used primarily for peer classroom observation.\textsuperscript{11}

The vast majority (10 out of 11) of our Big Ten respondents expect to have faculty reflect on their teaching and development of teaching effectiveness in narratives or through teaching portfolios. Student end-of-semester course ratings are used by all the responding institutions for formative assessment (i.e., improving teaching and self-evaluation) and for summative personnel actions of reappointment of non-tenure track faculty (faculty of the practice or clinical faculty), tenure decisions, promotion decisions, promotion of non-tenure track faculty (faculty of practice or clinical faculty). Typically the surveys are developed or authorized by the institution (Table 7) and used across departments, schools and the university (Table 8).

\textsuperscript{11} Further information on evaluation of teaching may be found in Appendices E (for Big Ten universities) and F (for other universities).
Appendix A: Overview of the teaching faculty at Rutgers University-New Brunswick

“Faculty” is one of those general occupational terms that presupposes an immediate understanding but quickly becomes muddled when attempts are made at providing definitional exactness. This is particularly the case when an effort is made to define the faculty role at a large, public research institution such as Rutgers-New Brunswick. The multiplicity of titles and ranks that faculty have, the many different functions and roles that faculty engage in, the different reporting requirements to federal, state, and higher education organizations, and the ever changing nature of higher education, with internal and external actors often altering expectations of faculty work, all contribute to a certain state of confusion about what constitutes faculty membership.

What follows is a brief description of the different ways one can be considered a faculty member at Rutgers-New Brunswick, including an additional look at the distribution of teaching among various instructional faculty groupings.

Roles and Titles

Perhaps the most widely held view of a faculty member is that of teacher - as the individual who stands in front of large (and small) classes of students imparting knowledge and wisdom to eager students. For many faculty at Rutgers-New Brunswick, this instructional role is an important component of who they are as a faculty member. Indeed, teaching is all that some faculty members do, while many others incorporate this instructional role with other activities that are associated with the university’s tripartite mission of teaching, research, and service. And just as some faculty only instruct, there are other faculty members that only participate in research or service activities.

While most faculty regardless of the particular activities they undertake (instruction, research, or service) work as full-time employees of the university, there is a substantial body of faculty who are hired as part-time employees. Full-time faculty are hired as either 9/10 month employees based on an academic year schedule or as 11/12 month employees who follow a calendar year schedule.

Another major distinction of faculty at Rutgers-New Brunswick is between faculty who are tenured or who are on the tenure track (often referred to as TT faculty) and other faculty who are hired as non-tenured or non-tenure track faculty (NTT faculty). Tenure is typically granted upon promotion to the rank of Associate Professor. Faculty members who have been granted tenure are considered to be appointed without term limitation and hold faculty membership indefinitely and at the “pleasure of the Board of Governors.”

Rutgers-New Brunswick faculty hold many different academic titles. Academic ranks for full-time faculty range from Assistant Instructor to Distinguished Professor. The highly diversified character of the university’s mission and organizational structure has resulted in many equivalent titles among faculty members. These ranks and equivalencies are duly noted and defined in Rutgers Policy 60.5.2, “Academic Titles and Rank Equivalencies for Tenured, Tenure-Track and
Non-Tenure Track Faculty.” The rank equivalencies listed in this policy implies equivalence in salary, tenure (where appropriate), academic rights and responsibilities, guarantees of academic freedom and due process, and comparable standards and procedures for appointment, reappointment, and promotion.

Other academic titles used at Rutgers to broadly characterize a faculty member include Professor Emeritus/a, University Professor, Visiting Professor, Part-Time Lecturer, Teaching Assistant, Graduate Assistant, Graduate Fellow, Post-Doctoral Fellow, and Post-Doctoral Associate. While individuals who hold these positions are not technically full participating members of the faculty, they do perform certain roles and functions often associated with faculty work.

Some faculty also may work in administrative positions at the university, such as members of the decanal leadership of a school or leaders in the university administration. During their time as administrators, these individuals are considered to be members of the university staff rather than the faculty, but at the conclusion of their administrative service they return as full members of the faculty.

Who Teaches and How Much do They Teach?

Teaching at Rutgers-New Brunswick is done by three main groups of faculty: tenured and tenure-track faculty, full-time non-tenured track faculty, and part-time non-tenure track faculty. TT faculty are mostly represented by those individuals who hold assistant, associate, and full professor (including distinguished professor) positions. NTT faculty mostly come from the ranks of assistant instructor and instructor positions, while part-time lecturers (PTLs) dominate the part-time contingent of faculty teaching, in conjunction with teaching and graduate assistants.

The first graph that follows shows the distribution of faculty at Rutgers-New Brunswick between Fall 2013 and Fall 2015. TT faculty who have instructional responsibilities comprise approximately 40 percent of all instructional faculty at Rutgers-New Brunswick, with full-time NTT faculty making up roughly ten to eleven percent. Part-time NTT faculty hold the largest percentage of the three faculty groups, approaching almost fifty percent of all instructional faculty at Rutgers.

The second graph shows this distribution of the three groups of instructional faculty for selected Rutgers-New Brunswick schools as of Fall 2015. Not surprisingly, the School of Arts and Sciences closely parallels the composition of instructional faculty for all of Rutgers-New Brunswick. The School of Environmental and Biological Sciences and the School of Engineering have higher percentages of their faculty employed as TT faculty, while many of the smaller academic units at Rutgers-New Brunswick have lower percentages of TT faculty compared to Rutgers-New Brunswick as a whole.

The distribution of teaching done by these New Brunswick faculty groups is depicted in the next two graphs. Using student credit hours (SCHs)\(^\text{12}\) as the metric for estimating instructional output, NTT faculty (both full-time and part-time) exceed the production of SCHs by TT faculty.

---

\(^{12}\) The metric of student credit hours is determined by the number of students taught in a course multiplied by the number of credits for that course, adjusted for shared teaching and mode of teaching.
by a ratio of more than two to one, with part-time NTT faculty accounting for nearly 40 percent of all teaching.

The great preponderance of instructional workload activity, as measured by the generation of SCHs, occurs at the undergraduate level, with over ninety percent of all instruction occurring at that level.\textsuperscript{13}

Because of its size, much like the distribution of faculty positions, the School of Arts and Sciences closely mimics the distribution of workload that was found for Rutgers-New Brunswick overall. SAS relies on part-time NTT faculty much more than either the School of Environmental and Biological Sciences or the combined total of all other New Brunswick schools.

\textsuperscript{13} Student credit hours is a popular and easily understood measure of workload, but in no way should it be taken to be the sole indicator of instructional productivity. Indeed, many different factors affect overall instructional activity, including activities spent outside of class time (office hours, course preparation, etc.), quality of instruction, and so on.
Distribution of Instructional Faculty
Fall 2013-Fall 2015

<table>
<thead>
<tr>
<th>Year</th>
<th>TT</th>
<th>FT Non-TT</th>
<th>PT Non-TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>43.2%</td>
<td>10.0%</td>
<td>46.8%</td>
</tr>
<tr>
<td>2014</td>
<td>41.3%</td>
<td>9.7%</td>
<td>49.0%</td>
</tr>
<tr>
<td>2015</td>
<td>39.8%</td>
<td>10.4%</td>
<td>49.8%</td>
</tr>
</tbody>
</table>

(N=3,074) (N=3,175) (N=3,279)
Distribution of Instructional Faculty by School
Fall 2015

<table>
<thead>
<tr>
<th>School</th>
<th>TT</th>
<th>FT Non-TT</th>
<th>PT Non-TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAS (N=1,490)</td>
<td>46.0%</td>
<td>20.5%</td>
<td>12.1%</td>
</tr>
<tr>
<td>MGSA (N=278)</td>
<td>68.7%</td>
<td>10.8%</td>
<td>11.2%</td>
</tr>
<tr>
<td>SEBS (N=277)</td>
<td>62.8%</td>
<td>26.0%</td>
<td>11.2%</td>
</tr>
<tr>
<td>ENG (N=175)</td>
<td>81.7%</td>
<td>7.4%</td>
<td>2.3%</td>
</tr>
<tr>
<td>SCI (N=175)</td>
<td>27.4%</td>
<td>38.0%</td>
<td>15.2%</td>
</tr>
<tr>
<td>BLOUSTEIN (N=79)</td>
<td>46.8%</td>
<td>70.7%</td>
<td>0%</td>
</tr>
<tr>
<td>ALL OTHER SCHOOLS (N=805)</td>
<td>20.7%</td>
<td>8.6%</td>
<td>70.7%</td>
</tr>
</tbody>
</table>
Distribution of Student Credit Hours (SCHs)

Total SCHs 455,202

TT 139,393 (30.6%)

Undergraduate 117,596 (25.8%)

Graduate 21,797 (4.8%)

Undergraduate 77,945 (17.1%)

Graduate 6,520 (1.4%)

Undergraduate 114,581 (25.2%)

Graduate 10,464 (2.3%)

FT Non-TT 134,786 (29.6%)

Undergraduate 128,267 (28.2%)

Graduate 6,520 (1.4%)

Undergraduate 50,322 (11.1%)

Graduate 6,520 (1.4%)

Undergraduate 55,978 (12.3%)

Graduate 10,464 (2.3%)

PT Non-TT 181,022 (39.8%)

Undergraduate 170,559 (37.5%)

Graduate 10,464 (2.3%)

Undergraduate 55,978 (12.3%)

Graduate 10,464 (2.3%)
School of Arts and Sciences
Distribution of Student Credit Hours (SCHs)

Total SCHs 293,035

- TT
  - Undergraduate 77,781 (26.5%)
  - Graduate 0 (0.0%)
- FT Non-TT
  - Undergraduate 85,250 (29.1%)
  - Graduate 0 (0.0%)
- PT Non-TT
  - Undergraduate 130,004 (44.4%)
  - Graduate 0 (0.0%)
School of Environmental and Biological Sciences
Distribution of Student Credit Hours (SCHs)

Total SCHs 18,640

TT
9,252 (49.6%)

Undergraduate
9,252 (49.6%)

Graduate
0 (0.0%)

FT Non-TT
7,497 (40.2%)

Undergraduate
7,497 (40.2%)

Graduate
0 (0.0%)

PT Non-TT
1,891 (10.1%)

Undergraduate
1,891 (10.1%)

Graduate
0 (0.0%)
All Other Schools (excludes SAS and SEBS)
Distribution of Student Credit Hours (SCHs)

Total SCHs 143,526

TT 52,360 (36.5%)
  - Undergraduate 30,562 (21.3%)
  - Graduate 21,797 (15.2%)

FT Non-TT 42,040 (29.3%)
  - Undergraduate 35,520 (24.7%)
  - Graduate 6,520 (4.5%)

PT Non-TT 49,127 (34.2%)
  - Undergraduate 38,664 (26.9%)
  - Graduate 10,464 (7.3%)
Notes:
Rutgers-New Brunswick excludes Pharmacy and Nursing. Undergraduate/graduate denotes level of course. Counts and percentages subject to slight rounding error. Based on Fall 2014.
Appendix B: Rutgers University Senate and New Brunswick Faculty Council
Charges, Reports and Recommendations on Evaluation of Teaching

Faculty Council Resolutions on the Evaluation of Teaching at Rutgers University
New Brunswick Faculty Council
Academic Regulations and Standards and Teaching Committee

Resolution on Evaluation of Teaching:

Whereas, the improvement of teaching and learning is a paramount goal of the University, and
Whereas, the evaluation of teaching is of great importance for achieving this goal, and
Whereas, there are wide variations in the way in which departments and programs carry out these
evaluations, and
Whereas, there is no clearly articulated university policy on how such evaluations should be carried out,
Therefore be it resolved that the NBFC recommends that the University require every department and
program to develop a detailed multifaceted plan for the evaluation of teaching, appropriate for their
organization and discipline, and
Be it further resolved that this plan shall require approval of the appropriate dean or director and that
these plans shall be available as public documents.

Approved January 26, 2007
http://nbfc.rutgers.edu/year06_07/ResolutionOnTeachingEvaluation_2007.html

Faculty Council Resolution on Evaluation of Teaching 2015-2016

RESOLUTION
The New Brunswick Faculty Council calls upon the Rutgers administration to work together with the
New Brunswick Faculty Council to create and charge a Task Force with developing a teaching evaluation
program that will, as accurately as possible, assess quality of instruction. The Task Force should be
broadly representative of the faculty and students. It should include tenure track faculty at various ranks,
nontenure track faculty, part time lecturers, and teaching assistants, as well as students from the sciences
and the nonsciences. It should also include academic administrators, a representative of CTAAR, and an
outside expert.

Rutgers must have a teaching evaluation program that will:

1. provide a rigorous and meaningful evaluation of the quality of faculty teaching for consideration in
decisions about the renewal, promotion, and tenure of faculty.

2. provide feedback on the quality of faculty teaching that will help units maintain their overall
commitment to pedagogy and help individual faculty members continue to develop their teaching skills.

In order to achieve these goals, the Task Force should address and remedy the defects of our current
Student Evaluation system; should rely on faculty to develop methods of evaluation that may often be
uniquely suited for their disciplines; and should be provided with such funding, staff support, and
professional expertise as the Task Force may require to fulfill its charge.
Charges to, Reports and Recommendations of University Senate on Evaluation of Teaching

University Senate Committee on Instruction, Curriculum and Advising
Review of Teaching Assistant Training
February 2001
http://senate.rutgers.edu/tatrain.html

Report of the Senate Faculty Affairs and Personnel Committee
On Charge S-0109
Best Practices in Assessment of Teaching
Paul Leath, Co-Chair
Barbara Lee, Co-Chair
February 12, 2002
As Revised May 3, 2002
http://senate.rutgers.edu/bestprac.html

University Senate
Faculty and Personnel Affairs Committee (FPAC)
Response to Charge S-1104 to the Instruction, Curricula, and Advising Committee (ICAC)
on Online Teaching Evaluations, and Best Practices in Evaluation of Teaching
Performance

University Senate Standing Committee Charges
Issued By The Senate Executive Committee, And
Pending As Of December 2016
http://senate.rutgers.edu/pendchgs.html
S-1510 - Instruction, Curricula and Advising Committee
Chair: Natalie Borisovets
Student Teaching Evaluations, and Best Practices in Evaluation of Teaching 2015: Revisit the March 2012 Instruction, Curricula and Advising Committee report on Charge S-1104, on Online Teaching Evaluations, and Best Practices in Evaluation of Teaching. Consider questions such as: Have there been changes in the completion rates and use of the Student Instructional Rating Surveys (SIRS)? How can we make students more aware that the results of the SIRS are available to them? Are the questions on the SIRS still valid, or do they need to be modified in light of changes in educational technologies? To what extent should the SIRS results be used in the evaluation of teaching and the promotion/tenure process? Are other Big-10 schools using such surveys, and to what purpose? Solicit input on personnel ramifications from the Senate’s Faculty and Personnel Affairs Committee. Respond to Senate Executive Committee by March 2017. [Issued November 2015.]
### APPENDIX C: Big Ten Evaluation of Teaching Survey Report

**Q16 - Please indicate your institution.**

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Illinois</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Indiana</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>IUPUI – will send document with required elements</td>
<td>0.00%</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Iowa</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Maryland</td>
<td>0.00%</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Michigan State</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Minnesota</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Nebraska – send letter (attached)</td>
<td>0.00%</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Northwestern</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Ohio State</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Penn State</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>Purdue</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>Rutgers</td>
<td>0.00%</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>University of Chicago</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>University of Michigan</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>Wisconsin</td>
<td>8.33%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
<td><strong>12</strong></td>
</tr>
</tbody>
</table>
Q2 - Do your Evaluation of Teaching practices include any of the following? (Please check all that apply.)

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>Answer %</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Peer review with class observations</td>
<td>83.33%</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Mentoring of new faculty in teaching</td>
<td>66.67%</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>Peer review of instructional materials</td>
<td>58.33%</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>Self-reflective statement</td>
<td>75.00%</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>Teaching Portfolio</td>
<td>41.67%</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>Focus-group interviews of students</td>
<td>41.67%</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>Exit interviews of graduating students</td>
<td>33.33%</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Alumni surveys</td>
<td>25.00%</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>End-of-course ratings surveys (i.e. instructor ratings and course review)</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>Use of written comment from end-of-course ratings surveys</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>Mid-course course survey</td>
<td>58.33%</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>Contributions to instruction</td>
<td>58.33%</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>Assessments of student learning</td>
<td>50.00%</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>Review of student work samples</td>
<td>33.33%</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>Scholarship of Teaching</td>
<td>66.67%</td>
<td>8</td>
</tr>
<tr>
<td>16</td>
<td>Videos of teaching</td>
<td>25.00%</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>Teaching awards</td>
<td>75.00%</td>
<td>9</td>
</tr>
<tr>
<td>18</td>
<td>Support of graduate student research</td>
<td>66.67%</td>
<td>8</td>
</tr>
<tr>
<td>19</td>
<td>Support of undergraduate student research</td>
<td>75.00%</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td>Employer ratings of student graduates for program evaluation</td>
<td>25.00%</td>
<td>3</td>
</tr>
<tr>
<td>21</td>
<td>Administrator ratings of teaching, scholarship and service</td>
<td>58.33%</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
<td><strong>12</strong></td>
</tr>
</tbody>
</table>
Q3 - Do you use any other assessments in the Evaluation of Teaching process? Please explain.

Some areas of the University also use decanal observation and mentoring of new faculty in teaching.

I checked all of the items above that I know are available to faculty members at the University of Minnesota. However, all of the items are not used by all faculty. And, I wouldn’t call some of these part of our evaluation of teaching (for example, I’m not sure how our support of student research helps us evaluate teaching -- though we have those types of programs). And, these processes are sometimes facilitated by the department, college, university, or our Center for Educational Innovation.

Most of the methods above are used in the evaluation of teaching for the process of tenure and promotion of assistant professors to associate professor. The evaluation of teaching is more rudimentary in the promotion process from associate to full, and for post-tenure review, and typically includes student evaluations, peer evaluations, review of instructional material, and in some instances scholarship on teaching.

This is a very difficult question to answer because the methods vary considerably from college to college and even between departments in a given college. For instance, some departments in our liberal arts college use peer observation, most do not. Similarly, one of our colleges solicits student letters, others do not. The items used universally are student ratings and a teaching statement.

Our P&T process requires: (1) Peer Review (course observation) (2) Student Feedback (a) Student Ratings of Teaching Effectiveness (Likert items) and (b) one other form of feedback from students. Typically, b) this is the written feedback from the open-ended questions that accompany the student ratings, but our procedures do permit collection of other kinds of student feedback, e.g. focus groups. It is up to the College to determine what the other form of student feedback is--so it could vary from college to college. However, I have never heard of any college doing anything other than using written feedback. And my office administers the SRs, so I’d probably know. Every other item on the list is used or done at Penn State, but is not necessarily part of the official evaluation for reappointment or promotion and tenure. If SoTL is considered, it is considered under Research section of the dossier, not the Teaching section.

http://www.psu.edu/dept/vprov/promotion.htm = University Guidelines Each college is allowed to specify some parts of the process; some colleges make their college P&T Guidelines public, others do not.

CLASSIs (Classroom Assessment by Student Interview) --although these fall until the mid-course assessment, they aren’t really surveys. Also, just to make clear--the Center for Teaching doesn’t conduct all of these. Depending on the assessment, the process might be carried out by an instructor, a chair, the Provost’s Office (e.g., teaching portfolio review), the Center for Teaching, etc.

Not all of the above are used university wide - in fact most occur in limited pockets, as every Tenure Initiating Unit (department or college) sets its own policy within guidelines. I selected those that I know of as being used either for summative review of individual instructors or for program level assessment of teaching effectiveness. One of our units also uses participation in a wide array of professional development activities with personal reflection on those chosen.

Self-review of things learned at development events attended
One thing that I want to make clear on the check-list above--not ALL departments do ALL of these things. Our procedures vary widely--wildly!--by department and school both in breadth and depth of evaluation of teaching. It also depends on whether the instructor is tenure line or teaching line. Teaching line faculty usually have more comprehensive measures of teaching performance, while tenure line focuses primarily on end-of-term student ratings, development of new courses, work with grad students and under grad research.
Q4 - Does your Evaluation of Teaching process take into account the diversity in instructional methods? (Please check all that apply.)

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lecture based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>Discussion based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>Case study courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>Small group interaction or seminar courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>Performance based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>Laboratory courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>Studio courses</td>
<td>100.00%</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>Fieldwork courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>Clinical work</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>Practicum</td>
<td>55.56%</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>Hybrid courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>Online courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>Lecture based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>14</td>
<td>Discussion based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>15</td>
<td>Case study courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td>Small group interaction or seminar courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>17</td>
<td>Performance based courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>18</td>
<td>Laboratory courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>19</td>
<td>Studio courses</td>
<td>100.00%</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td>Fieldwork courses</td>
<td>77.78%</td>
<td>7</td>
</tr>
<tr>
<td>21</td>
<td>Clinical work</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>22</td>
<td>Practicum</td>
<td>55.56%</td>
<td>5</td>
</tr>
<tr>
<td>23</td>
<td>Hybrid courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>24</td>
<td>Online courses</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100%</td>
<td>9</td>
</tr>
</tbody>
</table>
Q5 - If any of the above are checked, please explain.

I’m not sure what the question is asking specifically. If you are asking if we have specifically tailored and different evaluation processes for each of the types of courses, the answer is no. However, if you are asking if we evaluate each of the types of courses, the answer is yes. We recognize the diversity of methods, and -- particularly in evaluating faculty at transition points (promotion) -- departments are encouraged to evaluate them accordingly.

I’m not sure how to answer this. Are you asking whether there are specific methods targeted at each of these settings?

I’m not sure what you are asking. I presume that peer reviewers do not expect the same from faculty teaching 800 students as they do from faculty teaching a studio course, if that is so, then the answer is yes for all of these WRT Peer Review. The Student Ratings of Teaching Effectiveness has a departmental section, where the department chooses from a list of Faculty Senate approve questions and the department might choose to vary those questions by type of instruction, but they might not. However, there is no way to indicate the kind of course on our SR forms. The unit head writes a summary of each section of the Dossier (Teaching, Research, and Service) and this is where some department heads bring up such things as extraordinary. There is also a faculty statement, but it covers all three (T, R, S), and it is specifically not supposed to be self-evaluative (i.e. don’t tell your elders how you think you should be rated—they don’t care what you think!), it is primarily supposed to be descriptive, but I don’t think they’d be punished if it was somewhat self-reflective. Some departments do allow supplemental material to be included in the dossier, e.g. a teaching portfolio, but only at the department level. No supplemental material is seen by the college or university P&T Committees.

When invited to assess the teaching (and in some cases, the learning) in the types of courses indicated. We would use different approaches, depending on what the instructor is looking to find out and what kind of course (including size) is under consideration.

Methods used for these different instructional methods may vary, based on the type of course and the department.

Almost all of our units use a locally developed student feedback form for most didactic classroom and online experiences. Our health sciences colleges (Medicine, Health and Rehabilitation Sciences, Optometry, Dentistry, Nursing, Pharmacy, and Vet Med) all use a different instrument that was devised for clinical disciplines. Some of our Art and Music classes use an instrument developed for their specific context. Of course every unit uses different elements beyond the student survey, although most differ in ways that are not related to instructional methods, but more the culture of the discipline and department.

For our end of the term evaluation we have only two global questions. Then each department / school / college can have additional questions added to their standard "core". We do have departments which have created different cores for different types of courses such as labs, discussion, and the others checked above. On our online courses, we have some additional surveys that are sent out to gather more detailed information about performance of the student, rather than the performance of the instructor.

For online and blended courses, we have been using the Quality Matters rubric on campus for the last several years, and use is spreading. Regarding other teaching methods, the various colleges
and departments use a variety of tools and methods to evaluate teaching processes - there is a great deal of autonomy across the institution about how this happens.

I’m not completely sure I understand your question. The survey lists course types not methods...? So I believe that most departments will note differences in student ratings for different course types (e.g. she did great in seminars and small group sessions but she was perceived as teaching poorly in large lecture classes). Most of our schools use similar versions of our end of term student ratings, and that form does not distinguish among course types. Our medical school, however, has different types of ratings for different types of instruction.

There is no central evaluation of teaching at Purdue, only local. Since each of the instructional types above exist in the colleges, all contexts must be included.

Q6 - Does your Evaluation of Teaching process consider these additional components? (Please check all that apply.)

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Workload of classroom instruction (class size, course level, mode of instruction)</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Independent study courses</td>
<td>44.44%</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Team Teaching</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>Development of curricula, new courses and course materials</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>Teaching assistant supervision</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>Advising students in the major</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>Advising undeclared major students</td>
<td>33.33%</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>Workload of classroom instruction (class size, course level, mode of instruction)</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>Independent study courses</td>
<td>44.44%</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>Team Teaching</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>11</td>
<td>Development of curricula, new courses and course materials</td>
<td>88.89%</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>Teaching assistant supervision</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>13</td>
<td>Advising students in the major</td>
<td>66.67%</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>Advising undeclared major students</td>
<td>33.33%</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
<td><strong>9</strong></td>
</tr>
</tbody>
</table>
Q7 - If any of the above are checked, please explain.

If any of the above are checked, please explain.

This is highly variable depending on department/unit. Each of our units has a document guiding evaluation processes within that unit.

See above.

My sense is that faculty list these types of activities, although I cannot say how they are weighted in evaluation of teaching discussions. I am certain that the degree to which they are discussed would vary tremendously across the institution.

Advising is considered part of service. Independent study courses would be listed, but they do not count as part of faculty load. The department head might commend a faculty member for developing curricula, but it is unlikely to "count" much overall.

I don’t know how one would evaluate teaching/learning without considering the full array nature of the course.

Of faculty core dossier asks for a full listing of credit and non-credit teaching, including individual instruction and research supervision, as well as a discussion of course and curriculum development.

We do have end of semester evaluations for courses that have multiple instructors. What happens though is each instructor is evaluated separately. There can be questions added to the evaluation that references and asks about the team teaching practice.

Again, it varies from unit to unit.

Again, it varies by department and school to the extent that these "count." It also depends on the type of instructor (tenure line or teaching line).

Not sure how many are considered, since each area committee weighs elements differently.

Q8 - Do you use rubrics for any components of your Evaluation of Teaching process?

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>27.27%</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>72.73%</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100%</td>
<td>11</td>
</tr>
</tbody>
</table>
Q9 - If "Yes" would you be willing to share them? Would you please supply a link to the rubrics?

If "Yes" would you be willing to share them? Would you please supply a link...

We have both Student Ratings of Teaching (SRTs) which are standard questions, as well as resources for peer review of teaching. Link: http://www.academic.umn.edu/provost/peer_review/index.html

I don’t know, but probably not. Not unless they use a form we have created for peer review (e.g. http://www.schreyerinstitute.psu.edu/pdf/Classroom_Observation_Checklist_Form.doc). This is not really a rubric in the strictest sense, but it does result in some greater consistency across peer observations.

I would say that we don’t often use them unless we develop them for particular courses (and by request of the instructor). When we observe teaching, however, we might use a rubric.

some units might, but I have not seen them

Individual departments may use rubrics. I am only aware of some teaching observation/peer observation protocols.

Q10 - Do your Evaluation of Teaching practices allow for instructors to reflect on their development over time?

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>90.91%</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>9.09%</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100%</td>
<td>11</td>
</tr>
</tbody>
</table>
Q11 - If "Yes", please explain.

If "Yes", please explain.

<table>
<thead>
<tr>
<th>For every review process (and particularly at promotion and tenure time), instructors are asked to reflect on their teaching.</th>
</tr>
</thead>
<tbody>
<tr>
<td>At transition points, faculty are asked to submit a statement on teaching that takes account of development over time.</td>
</tr>
<tr>
<td>The teaching statement is the document that can be used for reflecting on development over time.</td>
</tr>
<tr>
<td>Not unless they are allowed to include philosophy statement or teaching portfolio as supplemental at the department-review level.</td>
</tr>
<tr>
<td>It can, if we evaluate them over time.</td>
</tr>
<tr>
<td>Faculty may make statements in dossiers that show growth over time. We encourage this reflection.</td>
</tr>
<tr>
<td>The core dossier also includes a Teaching Statement which can be used for this type of reflection.</td>
</tr>
<tr>
<td>Our instructors are given access to their end of semester scores. They can also request a longitudinal report of their performance on scores to match it with any changes in teaching methods they have used. This, along with support from our Instructor Development team, provides one method to improve teaching.</td>
</tr>
<tr>
<td>Tenure line faculty submit letters related to their reflection on their development; teaching line faculty are expected to reflect and improve/develop over time.</td>
</tr>
<tr>
<td>Instructors are free to pursue formative feedback and free to include whatever they like in their P &amp; T documentation.</td>
</tr>
</tbody>
</table>
Q12 - Are end-of-course surveys used in your institution or department for any of the following? (Please check all that apply.)

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Improving teaching</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>Self-evaluation</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>Reappointment of tenure-track faculty</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>Reappointment of non-tenure track faculty (faculty of the practice or clinical faculty)</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>Reappointment of casual or adjunct instructors</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>Reappointment of teaching assistants</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>Tenure decisions</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>Promotion decisions</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>Promotion of non-tenure track faculty (faculty of practice or clinical faculty)</td>
<td>100.00%</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>end-of-course surveys are not used for personnel decisions</td>
<td>0.00%</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>end-of course surveys are not used at this institution</td>
<td>0.00%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100%</td>
<td>12</td>
</tr>
</tbody>
</table>
The primary point of these evaluations is to improve teaching. If instructors are not inherently motivated, our use of these surveys for reappointment, tenure, and promotion helps motivate them.

The end-of-term ratings are used for a variety of purposes, both formative and summative. Formatively, individual faculty use the feedback to inform their future teaching. The ratings are also used as one component of summative evaluations at all levels.

Penn State is one of the few institutions at which the Student Ratings of Teaching Effectiveness were developed by the Faculty Senate for the purposes of tenure, promotion, and annual review. All contingent faculty (non-tenure track teaching faculty, full or part-time, and pre-tenure faculty) must have the SRTEs administered in every course. Each college has developed policies about the frequency of SRTE administration for tenured faculty. But students have come to expect to evaluate faculty in every course.

Any of the checked items can be used—and in some cases, are required to be used—as part of the portfolio of teaching submitted for tenure, promotion, awards.

The weight given to end-of-course surveys varies widely by department and school.

Student feedback, usually limited to end of term surveys, is a major portion of most evaluation of teaching. We urge a broader range of data.

The units make the decisions to the extent that they rely on end of course surveys for evaluation.

Our end of term ratings have multiple purposes: (a) faculty can use them to improve/learn about what worked in their classes etc; (b) administrators will use them for promotion and retention decisions and to benchmark across courses, sections, instructors etc; (c) students use them to tell other students about the class (they are online and visible to everyone at the institution)

Units use them differently.

Q14 - The end-of-course surveys were developed by.... (please check all that apply)

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>the Department</td>
<td>54.55%</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>the School</td>
<td>54.55%</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>the University</td>
<td>100.00%</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100%</td>
<td>11</td>
</tr>
</tbody>
</table>
Q15 - The end-of-course surveys are used by....(please check all that apply)

<table>
<thead>
<tr>
<th>#</th>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>the Department</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>the School</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>the University</td>
<td>91.67%</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>100%</td>
<td>12</td>
</tr>
</tbody>
</table>
APPENDIX D:

1) Evaluation of Teaching Rubric from Indiana University – Purdue University Indianapolis

2) Letter from University of Nebraska
# Documenting Teaching Performance

<table>
<thead>
<tr>
<th>Dimensions of Teaching Performance</th>
<th>Potential Locations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teaching Load</strong></td>
<td>List of courses, etc.</td>
</tr>
<tr>
<td><strong>Teaching Goals</strong></td>
<td>Goals and/or Teaching Philosophy</td>
</tr>
<tr>
<td><strong>Continuing Professional Development</strong></td>
<td>List of formal activities and their significance</td>
</tr>
<tr>
<td><strong>Use of Exemplary Teaching Methods</strong></td>
<td>Description of methods</td>
</tr>
<tr>
<td><strong>Quality of Teaching</strong></td>
<td>Reflective comments</td>
</tr>
<tr>
<td><strong>Evidence of Student Learning</strong></td>
<td>Reflective comments</td>
</tr>
<tr>
<td><strong>Ethics</strong></td>
<td>Self-report</td>
</tr>
<tr>
<td><strong>Scholarship of Teaching and National Leadership</strong></td>
<td>Publications, presentations, national leadership on teaching in discipline</td>
</tr>
<tr>
<td><strong>Course and Curriculum Development</strong></td>
<td>List of committees, etc.</td>
</tr>
<tr>
<td><strong>Recognition (Grants, Awards)</strong></td>
<td>List of recognitions</td>
</tr>
</tbody>
</table>

2015-16 Chief Academic Officer's Promotion and/or Tenure Guidelines
SUMMARY OF AREAS OF EXCELLENCE AND EXPECTATIONS FOR VARIOUS FACULTY CATEGORIES

<table>
<thead>
<tr>
<th>Advancement to</th>
<th>Area of Excellence(^1)</th>
<th>Other Areas of Performance</th>
<th>Expectation for External Peer Review of Case</th>
<th>Standard for Excellence (over and above record of quantity, quality, and impact of internal work)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Associate Professor Tenure Track</strong>(^2)</td>
<td>Teaching, Research and Creative Activity, or Professional Service</td>
<td>Satisfactory in areas not chosen for excellence as well as University Service as specified by the school. Highly satisfactory in all three areas for a balanced case.</td>
<td>Letters from independent(^6) peers, preferably in higher rank at peer or higher institution.</td>
<td>Record of nationally and/or internationally disseminated and peer-reviewed scholarship emerging national reputation.</td>
</tr>
<tr>
<td><strong>Professor Tenure Track</strong>(^2)</td>
<td>Teaching, Research and Creative Activity, or Professional Service</td>
<td>Satisfactory in areas not chosen for excellence as well as University Service as specified by the school. Highly satisfactory in all three areas for a balanced case.</td>
<td>Letters from independent(^6) peers, preferably in higher rank at peer or higher institution.</td>
<td>Record of nationally and/or internationally disseminated and peer-reviewed scholarship. A sustained national reputation as demonstrated by a well-established and cumulative body of work in rank. Special circumstances where scholarly productivity has been interrupted can be considered.</td>
</tr>
<tr>
<td><strong>Associate Librarian</strong>(^7)</td>
<td>Performance</td>
<td>Beyond satisfactory in either Professional Development or in Service and satisfactory in other area.</td>
<td>Letters from independent(^6) peers outside unit on IUPUI campus.</td>
<td>(No Additional requirements)</td>
</tr>
<tr>
<td><strong>Librarian</strong>(^8)</td>
<td>Performance</td>
<td>Excellence in either Professional Development or in Service and at least satisfactory in other area.</td>
<td>Letters from independent(^6) peers, preferably in higher rank at peer or higher institution.</td>
<td>Record of superior performance as an associate librarian and attainment of state, regional, or national recognition in the library profession (Indiana University Academic Handbook, UFC, 1978). Record of exceptional achievements in performance and a record of distinguished contributions to the university profession, or community in the secondary area of excellence. Quality is considered more important than mere quantity (Library Faculty Handbook, Promotion and Tenure Criteria for Librarians, 2004).</td>
</tr>
</tbody>
</table>

\(^1\) Balanced case expectations are defined by the Indiana University Academic Handbook as "balanced strengths that promise excellent overall performance of comparable benefit in excellence in one area and satisfactory in the other(s) to the University." This category applies to only tenure-track faculty and is to be used in exceptional cases.

\(^2\) For tenure decisions, tenure expectations are far performance commensurate with rank and evidence of continued service with distinction.

\(^3\) For tenure decisions, performance must be excellent, and professional development and service must be satisfactory. Tenure is granted to those librarians whose professional characteristics indicate they will continue to serve with distinction.

\(^4\) Balancing case exceptions for librarians only apply to the secondary criteria (to professional development, research and/or creativity and to service).

\(^5\) Independent is defined in the section on External Accomplishment.

\(^6\) For more detailed information regarding evaluating librarian performance, please review the "Suggested Standards for Evaluating Librarian Performance."
Dear Dr. Devanas,

I’m not sure what you’re looking for on this survey. At Nebraska, we require a peer evaluation of teaching at the time of promotion and/or tenure consideration, but this peer evaluation can be performed in different ways depending on the college or even on the department within a given college. Each of the Evaluation of Teaching practices you list is used by one or more units on our campus, but I’m not sure that any of them are completely universal. Given the lack of a central "evaluation of teaching policy", I don’t know how to answer the rest of the questions in a way that would be useful to you.

I hope this brief email helps. Please feel free to contact me if I can provide further information.

Best regards, Judy Walker

--------
Judy Walker  Aaron Douglas Professor of Mathematics  Interim Associate Vice Chancellor for
Faculty Affairs
University of Nebraska -- Lincoln
judy.walker@unl.edu
APPENDIX E: Links to Big Ten Evaluation of Teaching

University of Illinois
Role of teaching in promotion and tenure: See Section 4.
https://faculty.uic.edu/promotionandtenure/
Student Surveys:
https://citl.illinois.edu/citl-101/measurement-evaluation/teaching-evaluation/teaching-evaluations-(ices)

Indiana University
Role of teaching in promotion and tenure: See Section 7.
Student Surveys:
https://one.iu.edu/task/iub/course-evaluations

Indiana University – Purdue University Indianapolis
Promotion and Tenure Guidelines: documenting teaching page 17
https://academicaffairs.iupui.edu/PromotionTenure/IUPUI-Guidelines/Resources
Student Surveys:
http://ctl.iupui.edu/Resources/Documenting-Your-Teaching/Assess-Teaching-Effectiveness

University of Iowa
Faculty review, promotion, and tenure: section viii for teaching materials
https://provost.uiowa.edu/faculty-review-promotion-tenure
Student Surveys:
https://teach.its.uiowa.edu/technology-tools/ace-online-course-evaluations

University of Maryland
Appointment, promotion and tenure of faculty: page 16 details on teaching
https://president.umd.edu/administration/policies/section-ii-faculty/ii-100a
Student Surveys:
https://courseevalum.umd.edu

Michigan State University
Faculty guide for reappointment, promotion and tenure review
https://www.hr.msu.edu/promotion/facadstaff/FacGuideTenure.htm
Student Surveys:
https://login.msu.edu/?App=J4300

University of Michigan
Faculty Promotion Guidelines:
https://www.provost.umich.edu/faculty/promotion_guidelines/procedures.html
Student Surveys:
http://ro.umich.edu/evals/
http://www.crlt.umich.edu/node/57837

University of Minnesota
Promotion and tenure: see section 7.11
http://www.academic.umn.edu/provost/faculty/tenure/
Student Surveys:
https://oms.umn.edu/srt/
University of Nebraska
Promotion and Tenure
http://academicaffairs.unl.edu/faculty/promotion-tenure
Assessing Teaching
http://www.unl.edu/gtahandbook/assessing-your-teaching

Northwestern University
Role of teaching in promotion and tenure:
Student Surveys:
http://www.northwestern.edu/ctec/

Ohio State University
Policies and Procedures:
https://oaa.osu.edu/policiesprocedureshandbook.html
Student Surveys:
https://registrar.osu.edu/faculty/sei/sei.asp

Pennsylvania State University
Role of teaching in promotion and tenure: see page 6
http://vpaa.psu.edu/files/2016/09/p_and_t_faq-1gi48nm.pdf
Student Surveys:
https://www.srte.psu.edu

Purdue University
Role of teaching in promotion and tenure: see page 6
Student Surveys:
https://www.purdue.edu/cie/Website%20CoursEval/courseeval/

University of Chicago
Appointment and Promotion
https://provost.uchicago.edu/handbook/academic-appointments/appointment-and-promotion
Student Surveys:
https://shibboleth2.uchicago.edu/idp/profile/SAML2/Redirect/SSO?execution=e1s1

University of Wisconsin
Role of teaching in promotion and tenure: see page 2
http://www.seefac.wisc.edu/newfac/tenure/tenure_process.pdf
Student Surveys:
https://testing.wisc.edu/standardizedcourseevals.html
Appendix F: Links to other universities

**Stanford University**
Promotion to Tenure: see section 3.7
https://facultyaffairs-humsci.stanford.edu/chapter-3-promotion-tenure-tenure-line
Student Surveys:
https://vptl.stanford.edu/faculty-instructors/evaluation-feedback/stanfords-course-evaluations

**Harvard University**
Promotion handbook:
http://academic-appointments.fas.harvard.edu/internal-promotion-tenured-professor-tenure-track-position
Student surveys
http://registrar.fas.harvard.edu/courses-exams/course-evaluations

**Princeton University**
Dean of the Faculty information
Student Ratings – need log in
Chapter 3: Suggested frameworks for evaluation of teaching at Rutgers University-New Brunswick

In this chapter, we discuss frameworks for the evaluation of teaching at Rutgers University-New Brunswick (RU-NB). We use the word “frameworks” in the plural to emphasize that no single framework for evaluation of teaching can reasonably be applied to all faculty. Accordingly, we first set out some general principles for evaluation of teaching, but then apply these – in somewhat different ways – to evaluation of the three main kinds of faculty at RU-NB: tenured and tenure-track (TT) faculty; full-time non-tenure-track (NTT) faculty; and part-time lecturers (PTLs). We conclude with a brief discussion of evaluation of on-line teaching.

A. Guiding principles for evaluation of teaching at Rutgers University-New Brunswick

We come to this document with a set of assumptions that will have to be aligned with the introduction. Fundamental ideas include:

- Evaluation of teaching should never be dependent on one measure
- The wide variety of course experiences offered across and within the many disciplines within RU-NB mandates flexibility in the evaluation process. At the same time, certain standards of evaluation must exist for the process to be useful, especially for the reappointment and promotion process.
- Within the RU-NB environment, the structure of any evaluation system will include a great deal of local decision-making and choice.
- Evaluations are a process that should provide valuable information to faculty as well as to departments about the quality of teaching that students are experiencing.
- Tenure-track, non-tenure track, and part-time faculty should all participate in the evaluation system though timing, frequency, and purpose of evaluation events may vary.
- To the extent that social processes involving multiple individuals are brought into any evaluation, there is a greater likelihood that shared understandings of quality teaching will be obtained.
- Creativity, innovation, and risk-taking are important to the development of excellence in teaching and the evaluation process should reflect their value.

Considerations related to faculty type and rank

As stated above, we believe the evaluation of teaching should be carried out for all faculty; tenure-track/tenured, non-tenure-track, and part-time lecturers. However, some aspects of the evaluation process will differ depending on the type and/or rank of the faculty member. In order to ensure the evaluation process is carried out efficiently, effectively, and appropriately for all faculty members we suggest the following considerations pertaining to the purpose, timing, and frequency of the evaluation process be kept in mind.

For TT and NTT faculty and their departments, the evaluation process could serve several potential purposes including, but not limited to, the following:
1. Providing information for reappointment, promotion, and teaching award decision-makers (including Form 1A for TTs and Evaluation Short form for NTTs);
2. Supporting faculty mentoring;
3. Providing feedback for faculty professional development.

The frequency and timing of evaluation would depend, in part, on its purpose. We suggest that for TT faculty, the first evaluation be carried out in sufficient time to be included in the pre-tenure review process and again for subsequent promotion reviews, or at least once every 4 years. Once tenured, evaluations should continue to be carried out for inclusion in post-tenure reviews. The timing and frequency of evaluation of NTT faculty would similarly be tied to the faculty member’s reappointment/promotion schedule. Departments using the evaluation process as part of their faculty mentoring process may wish to carry out the teaching evaluation earlier in the faculty member’s career and/or more frequently.

Unlike tenure-track/tenured and non-tenure-track faculty, PTLs are not hired under long-term contracts and typically are not eligible for promotion. PTLs also enjoy far less integration into the culture and practices of their respective departments, which can make the desired goals of a particular course less clear. We suggest that departments keep this in mind when designing and implementing the evaluation of their PTL faculty.

For example, given that there is no University-wide mechanism or structure for PTL promotion, the purpose of the evaluation relative to a PTLs continued employment would likely be limited to assessing whether the PTL is meeting expectations; questions regarding patterns of growth or development in teaching may be less germane. Departments may therefore wish to limit the frequency and/or complexity of the evaluation of their PTLs relative to their TT and NTT faculty.

Departments may find that in addition to providing feedback regarding the PTL’s teaching effectiveness for both the department and the PTL, the evaluation process may serve as a mechanism by which the department’s expectations are conveyed to their part-time faculty. These expectations may include information specific to the course being taught (e.g., content, depth) as well as the general academic culture and expectations of the department (e.g., appropriate assignments and assessments, grading policies, feedback to be provided to students, etc.)

However departments decide to carry out evaluations of the tenure-track/tenured, non-tenure-track and PTL faculty, we suggest they do so with the goal of continued improvement in classroom instruction foremost in mind.
B. **Possible evaluation methods: Introduction**

In this section, we offer a range of possible evaluation methods (and associated policies) for departments to consider. We suggest that every evaluation system should include evidence from three types of sources that each support different inferences about teaching quality:

A. Philosophy and planning documents  
B. Evidence from classroom practice  
C. Student feedback

Furthermore, given the immense diversity of courses and classroom experiences to be evaluated, we suggest departments consider the following as they design and carry out their evaluations, as they may influence the appropriateness of specific kinds of evidence as well as expected outcomes:

- Class size / type (large lecture, online, seminar, active learning, etc.)  
- Course level (introductory, senior capstone, etc.)  
- Is the course required? Does the course have popular interest?  
- Is the course open to majors only or to the general population?  
- Does the course have a quantitative orientation? Is it reading or writing intensive?

The following sections elaborate on these evidence sources, including ideas on how to evaluate each type of evidence in order to make judgments of quality and provide meaningful feedback. We include some candidate criteria that are not exhaustive but are intended to provide some starting points for departmental discussions.

**Philosophy and planning documents**

This class of evidence includes teaching statements and syllabi. Both types of documents can provide insight into conceptions of teaching and content, organization, clarity, and planning. The teaching statement will likely convey higher-level and broader perspectives on teaching than will a syllabus, which will provide more detailed information about a particular course.

If a department chooses to use either or both types of evidence, we recommend that the department set up two evaluative supports:

1. An evaluative rubric  
2. A repository of strong exemplars

**Syllabi** - We suggest that a faculty group establish a set of criteria that contribute to an effective syllabus. This can be done by sampling a set of departmental syllabi and coming to some consensus about the attributes of strong syllabi. These criteria should then be codified and
shared with all faculty. Criteria should go beyond the formal syntax required in Rutgers syllabi and address substantive and stylistic issues as well. Current syllabi can be reviewed and exemplars selected to provide illustrations of strong syllabi and how they satisfy the defined criteria. These should be posted on a departmental resource (e.g., website) available to all faculty.

For evaluation, one or more faculty members can review the syllabus with the faculty member being evaluated using the rubric to guide the discussion. We suggest that syllabi from previous semesters be included in this review in order to provide evidence of professional growth and innovation, etc.

Potential elements to consider for inclusion in the syllabus rubric:

a. *Basics* - to what extent does the syllabus include information such as period and classroom, contact information for instructor (and TA), schedule (including dates for major assignments/exams? Is the information presented clearly, logically?

b. *Learning objectives* - to what extent are the objectives of the course presented clearly and concisely? Do the assignments align with the learning goals, and is this alignment presented clearly?

c. *Content* - to what extent is the content appropriate for the course and at the appropriate level of depth? Are critical concepts addressed?

d. *Organization* - to what extent do readings and instructional plans support course objectives?

**Teaching Statements** - We suggest an analogous process for the evaluation of teaching statements. This can be done by sampling a set of departmental teaching statements and coming to some consensus about the attributes of strong teaching statements. These criteria should then be codified and shared with all faculty. Criteria should address substantive and stylistic issues.

Current teaching statements can be reviewed and exemplars selected to provide illustrations of strong teaching statements and how they satisfy the defined criteria. These should be posted on a departmental resource (e.g., website) available to all faculty.

For evaluation, one or more faculty members can review the teaching statements with the faculty member being evaluated using the rubric to guide the discussion.

Elements in the teaching statement rubric might include:

a. *Teaching Philosophy and Goals* - to what extent are the philosophy and goals thoughtful and coherent.

b. *Clarity* - to what extent are learning goals and expectations clearly stated?

c. *Evidence* - to what extent are specific examples and personal experiences provided,
rather than a reliance on educational buzzwords and generalities?

d. **Assessments** - to what extent are appropriate and practical assessment methods included? Are the assessments aligned with the stated goals?

e. **Student-focused** - to what extent is there evidence of a focus on student learning rather than simply content delivery? To what extent is there evidence of sufficient knowledge and appreciation of the diversity of backgrounds, learning abilities, interest levels, etc., of the student body?

f. **Self-reflection and pedagogical development** - to what extent is there evidence of thoughtful reflection on past experiences and challenges and plans for further development?

Further information regarding creating and evaluating syllabi and teaching statements is available from the Rutgers Center for Teaching Advancement & Assessment Research (https://ctaar.rutgers.edu/teaching/). Additionally, we have collected several resources from other universities that departments might find useful, including syllabus rubrics from the University of Virginia Teaching Resource Center (Appendix A) and Cornell University Center for Teaching Excellence (Appendix B) and a discussion on crafting and assessing teaching philosophy statements (Appendix C) from The Center for Research on Learning and Teaching at The University of Michigan.

**Evidence from classroom practice**

This class of evidence includes classroom observations and classroom artifacts (e.g. exams, assignments, lab work, student papers). Each of these can provide important information about expectations for students in terms of content, reasoning and problem-solving, pedagogical strategies, and the quality of work being produced in classrooms by students.

**Observations** - We suggest that faculty develop or adopt/adapt a classroom observation protocol, with particular focus on areas that represent the best opportunity to improve instruction within the department (as well as for particular faculty). A protocol can at least provide some common structure and language across observations within a department.

In observing, it is critical to focus on both the teacher and the students in the class: evidence that they are engaged intellectually is perhaps the most important area of focus. Departments will have to decide how they evaluate both the quality of content and the pedagogical skills in teaching that content within an observation structure.

Elements in the Class Observation protocol/rubric might include:

a. **Intellectual Engagement** - To what extent are students in the class intellectually engaged in the course content? This can be reflected in student participation, non-verbal cues etc.
b. *Discourse* - To what extent is there an exchange of ideas between and among students? To what extent is the teacher asking questions that spur thinking and discussion?

c. *Content* - To what extent is relevant content introduced clearly and accurately? To what extent are explanations, representations, models, and metaphors (etc.) supportive of student learning?

d. *Instructor Presence* - To what extent does the instructor exhibit a passion for the material being presented and/or a commitment to ensuring the students “get it?” This element may include intangibles (“high energy,” “passionate”) as well as criteria such as timeliness (begins and ends on time), making eye-contact and creating a good rapport with students.

Alternatively, classroom observations could be based on the model proposed by Chickering and Gamson, entitled "Seven Principles of Good Practice in Undergraduate Education," published in 1987 but still relevant today.¹ According to Chickering and Gamson, good practice in undergraduate education includes the following:

1. Encourages contact between students and faculty, implying faculty should make themselves available and accessible to students, so as to answer questions or concerns.

2. Develops reciprocity and cooperation among students, implying that good learning is collaborative and social, as this helps in the meaningful acquisition of skills and knowledge.

3. Encourages active learning, implying that students are not empty vessels to be filled with information, but they are most likely to enhance their own learning and accomplish the course goals if they are active participants.

4. Gives prompt feedback, as this can help students reflect on what they have already learned and what they still need to know.

5. Emphasizes time on task, implying that the teacher knows how to optimally pace the class, so that students are neither overwhelmed by the pace being too fast, nor are bored if it is too slow.

6. Communicates high expectations, which should be communicated explicitly on the syllabus at the start of the semester, as knowledge of the teacher’s expectations can help students to strive towards academic excellence.

¹ See http://www.lonestar.edu/multimedia/sevenprinciples.pdf.
7. Respects diverse talents and ways of learning, which is particularly important at Rutgers because of the diversity of student backgrounds, and is important in general because of the multiple intelligences students might bring to the classroom.

Departments will also have to judge who and how many observers are used. To the extent that classes are video-recorded, the lesson can be viewed multiple times, including by the instructor and a mentor. There are multiple systems available for the storage of videos within observation contexts which might prove to be useful.

Research is clear that the characteristics and quality of teaching can vary substantially across lessons. Therefore, while feedback can be provided for each individual lesson, any summative judgment of a faculty member, particularly for tenure and promotion, should be based on multiple observations.

Classroom Artifacts - Examinations of the work that students are asked to do can be very revealing about the intellectual demands on students and how they respond to those demands evident in the work they produce. Departments should make decisions about the following when they evaluate a faculty member:

1. What artifacts should be collected (assessments, research papers, labs, grading/feedback to students, etc.)? How much and what range of student work should be collected? Typically, it is appropriate to collect multiple and varied samples of work to get a more representative perspective on what students are doing.

2. How will the evidence be judged? As with other sources of evidence, a faculty committee should develop a rubric that articulates the criteria for effective artifacts of one type or another.

Elements in the Classroom Artifacts rubric might include:

a. Analysis - to what extent are students being asked to reason about important ideas in a specific discipline (in contrast with recall/procedural execution)

b. Communication - to what extent are students being asked to communicate their ideas through explanation, rich description, elaborated writing, effective representations (appropriate to a discipline)

c. Alignment - to what extent are the activities/assessments aligned with the objectives or goals of the course

Student feedback

Student feedback on courses can be elicited through various means:
Surveys: The SIRS system is already in place, and we recommend its continued use as it is a fully operating system that provides useful information when placed in a broader context of other information. The SIRS data should not be used as the sole source of information about a faculty member’s teaching, however. (For much more discussion of student instructional ratings surveys, see Chapter 4.)

While individual faculty members can currently add questions of their own, the department faculty should decide whether there are any additional questions they would like to add for all (or specified parts) of the department faculty.

Student focus groups. Focus groups have the potential to yield important information about individual faculty member or collective faculty (e.g., program). Departments will have to make decisions about which faculty are present, who facilitates, how and how many students are selected etc. Focus groups are likely to be more effective if they follow some type of structure and model best practices for such a process. Departments will have to set policy on the recording and synthesis of focus group discussion. Are focus groups recorded? If yes, are recordings transcribed? Who synthesizes and reports out the takeaways? Is there a debriefing process with the faculty instructor?

Exit Interviews. Interviews with or a survey of graduating majors/minors could reveal information regarding the lasting impact a faculty member may have had on his/her students. Possible topics for discussion might include, e.g., The most useful (or memorable, or life-changing...) course/instructor; The instructor who has had the biggest impact on me; The paper/project I’m most proud of; What I remember most from instructor X’s course(s).

Portfolios

Any or all of the sources of evidence described above could be included in a teaching portfolio. While each of the pieces can still be evaluated separately, there would also be a process by which information is evaluated more holistically. Presumably, there would be significant interaction around the portfolio between the faculty instructor and the faculty mentor(s). If pursuing a portfolio, decisions need to be made about the requisite evidence, including how much written commentary the faculty instructor needs to provide.2

C. Part-time lecturers and the evaluation of teaching

At present, there is no university-wide policy requiring the evaluation of part-time lecturers at RU-NB. As noted in Chapter 2 of this report, some programs (Expository Writing, for example) have robust procedures for evaluation of their NTT and PTL faculty. However, in

---

2 An excellent resource for assembling and assessing teaching portfolios is available from The Center for Research on Learning and Teaching at The University of Michigan (see Appendix D).
the majority of cases, PTLs do not undergo any formal evaluation procedures at any point in their teaching careers.

Many, if not most, PTLs would be very open to the idea of evaluation and/or mentoring by senior faculty members or other qualified department personnel. Having a clear idea of what is expected of the instructor in the classroom would be seen by most PTLs as a welcome improvement over existing conditions; a senior faculty mentor could provide such orientation. Additionally, announced visits to the classroom by one or more of the departmental faculty could also help the PTL gain a better understanding of the expectations and requirements of the department. These steps could be supplemented by an examination of class assignments by the senior faculty to further determine the effectiveness of the instructor and to recommend ways to make progress in learning goals. Subject to the considerations noted in Chapter 4 (see below), a component of this effort could include results of student instructional ratings surveys such as SIRS. By these means, the department and the PTL would obtain a balanced picture of the PTL's performance in the classroom and be able to identify ways in which the instructor could improve. The ultimate benefit of such a commitment to systematic, comprehensive evaluation of PTL teaching should be continued improvement in classroom instruction, a key factor in improving student academic performance at Rutgers.

Although many PTLs would find evaluation of classroom teaching to be beneficial, it may also create special problems for them, because the teaching experience of the typical PTL is quite different in many respects from that of full-time faculty (either NTT or TT). PTLs enjoy none of the protections afforded to their full-time counterparts; they have no long-term contracts; in virtually all cases, they are appointed on a semester-to-semester basis with no guarantee of reappointment. In this setting, the costs and benefits of teaching evaluation would be quite asymmetric for most PTLs. In practice, the best possible outcome of a favorable teaching evaluation for the PTL would simply be reappointment to his or her existing position, with no possibility of advancement of any sort. On the other hand, an unfavorable teaching evaluation could lead to the PTL's dismissal or non-renewal.

Likewise, other aspects of the typical PTL's situation make teaching evaluation of PTLs more problematic than evaluation of other (NTT or TT) faculty. For example, relative to other faculty, PTLs are more likely to teach large-enrollment survey courses, in which it is harder for the instructor to engage with students in a meaningful way – a limitation that may translate into lower student instructional ratings. PTLs also enjoy far less integration into the culture and practices of their respective departments, often making the desired goals of a particular course less clear.

Further, since PTLs customarily do not receive long-term appointments, it becomes more difficult to say when their departments would conduct evaluations of those PTLs. It would be unworkable, and probably unnecessary, to evaluate PTLs every semester; evaluation after some predetermined number of semesters would be more workable. But how frequent should such
evaluations be? Individual departments will need to strike a balance between the costs (in terms of time and effort) and added benefit (in terms of more up-to-date information) of more frequent evaluation. In the case of full-time faculty, there are several decision points at which it would be natural and appropriate to conduct evaluations of teaching: decisions on reappointment, tenure, promotion, post-tenure review, etc. In contrast, PTLs are contingent faculty, with no clear path to advancement within their departments.

Finally, introducing teaching evaluations for PTLs will mean requiring PTLs to undertake additional uncompensated work for use in the evaluations (e.g., prepare teaching portfolios and statements). If the only reward for this additional work would be the opportunity for the PTL to retain the position that he or she currently holds, it is not hard to see why most PTLs would be skeptical about the benefits of teaching evaluation.

We are aware that a number of universities in the Big Ten\(^3\) and elsewhere have developed policies and procedures for professionalizing their contingent-faculty workforce (e.g., PTLs and similar positions), and that these procedures usually provide for teaching evaluation. It may be that developments of this kind at RU-NB can provide a means of introducing teaching evaluation for PTLs as well as other improvements in their status. Further discussion of this topic would take us far beyond our terms of reference. We therefore content ourselves here with simply noting that the question of teaching evaluation for PTLs has potentially important implications for professionalizing the PTL work force, and also vice-versa.

D. **Online courses and the evaluation of teaching**

The number of online courses is growing quickly throughout the university, for all ranks of faculty, from PTLs to tenured professors and all positions in between. Such courses are often considered to be directly equivalent, in terms of credit-hours, to their counterpart courses offered on campus.

Evaluating success in an online course may well be difficult, perhaps even more difficult than evaluating courses given on campus. An online course replete with PowerPoint slides, fancy animations and other flashy displays may be worthwhile, but, as with an on-campus course, these features do not necessarily guarantee success. Student instructional ratings may be useful in assessing online courses, but as noted in Chapter 4, the information provided by such ratings may well be ambiguous or misleading. Similarly, students' grades may not reveal much about the quality of an online course. Should we look at the attrition rate? Without information on why students drop a course, it will not be possible to make meaningful inferences: some students who drop a course might merely have been shopping for courses, or might have dropped for reasons entirely unrelated to the course itself. Should student survey ratings, grades and

\(^3\) See Appendix E, which outlines a schedule by which part-time faculty at the University of Michigan are evaluated.
attrition rates in online courses be compared with data for students taking the equivalent on-campus? It seems that this sort of comparison between online and traditional courses makes no sense, as teaching and learning in these two environments are likely to be quite different.

One helpful method by which to assess an online course might be to begin with the model called “Seven Principles of Good Practice in Undergraduate Education” (already mentioned in this chapter), which has been specifically adapted for online classes by Chickering and Ehrmann. They believed that “[i]f the powers of the new technologies is to be fully realized, they should be employed in ways consistent with the Seven Principles” (see Chickering and Ehrmann, “Implementing the Seven Principles: Technology as Lever,” AAHE Bulletin, 1996). It would also seem that many of the same tools and techniques discussed earlier in this chapter (e.g., use of teaching portfolios) can be adapted to evaluation of online courses.

To sum up, as the number of online courses continues to grow, the University will need to develop a systematic set of procedures for evaluating online courses, keeping in mind the ways in which such courses differ from, and are similar to, on-campus courses.
Appendix A
Measuring the Promise: A Valid and Reliable Syllabus Rubric

Guide to Assessing the Focus of Syllabi

Michael Palmer, Dorothe Bach, & Adriana Streifer
University of Virginia, Teaching Resource Center

Contents:
- Overview
- Rubric
- Validity
- Scoring
- Supplemental Rubric
- References
- Appendix A: Verbs for Significant Learning
- Appendix C: Reference Syllabi (available as a separate document)
- Appendix D: Scored and Annotated Reference Syllabi (available as a separate document)
Overview

This rubric was designed to help quantitatively and qualitatively assess the degree to which a syllabus achieves a learning-centered orientation. The development of the rubric was guided by the literature on learning-focused course design (Fink, 2013; Hansen, 2011; Wiggins & McTighe, 2005; Wulff & Jacobson, 2005), teaching (Ambrose, Bridges, DiPietro, Lovett, & Norman, 2010; Biggs & Tang, 2007; Blumberg, 2008; Nilson, 2010), and student motivation (Schunk, et al., 2007; Svinicki, 2004). The rubric design was also influenced by existing literature on syllabus construction and syllabus components (Baecker, 1998; Becker & Calhoon, 1999; Canada, 2013; Doolittle & Siudzinski, 2010; Eberly, Newton, & Wiggins, 2001; Garavalia, Hummel, Wiley, & Huitr, 1999; Habanek, 2005; Harnish & Bridges, 2011; Matejka & Kurke, 1994; O’Brien, Millis, & Cohen, 2008; Parkes, Fix, & Harris, 2003; Parkes & Harris, 2002; Singham, 2007; Slattery & Carlson, 2005; Smith & Razouk, 1993). It accounts for nuances in syllabi while also maintaining widespread relevance to courses in a diverse range of disciplines, levels, and institutions.

The rubric focuses on four criteria typical of learning-centered syllabi: (1) learning goals and objectives, (2) assessment activities, (3) schedule, and (4) overall learning environment. These criteria do not map onto any specific section of a syllabus (with the exception of the Schedule); instead, users are directed to search for evidence of the quality of all criteria across the syllabus.

We break down each criterion into multiple components, and provide a range of options for what evidence of proficiency in those components might look like. For example, the criterion of Overall Learning Environment contains components such as positive tone, fostering motivation, and high expectations, each of which syllabi may signal in a variety of ways, from giving students a degree of control over their learning experience, to offering resources to help them succeed, to opening with enthusiastic language that communicates the opportunity to wonder and ask questions about the course material without fear of criticism.

Each of the 16 components—designated as essential, important, or less-important—is scored on the strength of supporting evidence. Strong evidence indicates that many (but not necessarily all) of the characteristics of the component are present in the syllabus and match the descriptions closely. Moderate evidence indicates that a few of the characteristics of the component are present in the syllabus and/or only partly match the descriptions. Low evidence indicates that very few of the characteristics of the component are present in the syllabus and/or don’t match the descriptions.

You may use our syllabus rubric for research purposes as long as you provide reference to the following:


For the norming process, we recommend users score the reference syllabi first without the aid of our scores and annotations. Then, compare scores, reading the annotations when discrepancies exist.

Questions? Contact Michael Palmer at mpalmer@virginia.edu.
Rubric

“Essential” components are shown in gold, “important” components in silver, and “less-important” components are in white.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>What the component looks like:</th>
<th>Ideas for where to look and examples of what to look for (not all need to be present):</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning Goals &amp; Objectives</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Learning goals and objectives are not an “afterthought,” but are a central element of the course.</strong></td>
<td></td>
</tr>
</tbody>
</table>
| 1.                                 | Explicitly or implicitly **stated learning goals** (i.e. long-range, high-level goals) encompass the full range of Fink’s dimensions of significant learning (i.e. knowledge, application, integration, human dimension, caring, learning how to learn). | • Learning goals are often found in the course description, especially affective goals.  
• Implicit goals may appear in other sections of the syllabus (e.g., assessment, schedule, tips for student success). |
| 2.                                 | **Course-level learning objectives** are clearly articulated and use specific action verbs to describe in measurable terms what students will be able to do, value, or know at the end of the course. Like the goals they are derived from, the learning objectives map onto the full range of Fink’s taxonomy. | • Course-level learning objectives are in a prominent and easily identifiable location (i.e., labeled section).  
• Learning objectives with non-quantifiable terms, such as “understand” and “know,” are avoided. For examples of strong verbs, see Appendix A: Verbs for Significant Learning.  
• The syllabus considers the full range of Fink’s taxonomic dimensions (including the affective ones). It is not necessary that course objectives list affective dimensions, as long as there is evidence elsewhere in the syllabus that they are being considered.  
• Typically, 5–8 course-level objectives are appropriate. More or fewer could be problematic. |
| 3.                                 | **Learning objectives are appropriately pitched** to the course level, class size, position of the course within the curriculum, and characteristics of students taking the class. | • This is likely difficult to assess without knowledge of the discipline and curriculum. |
| **Assessment Activities**          |                                                                                                 |                                                                                                                                                      |
|                                    | **All major assessment activities positively support the learning objectives.**                                                                 |                                                                                                                                                      |
| 4.                                 | It is clear that the **objectives and assessments are aligned**. In other words, the major assessment activities map onto the full range of learning objectives and the degree of mapping correlates with the weighting of the assignment. | • Though a complete mapping may not be possible without input from the instructor, connections between the objectives and major assessments should exist (i.e., the each major assessment activity should map to one or more learning objectives). |
| 5.                                 | The basic features of the **major summative assessment activities are clearly defined**. The assessment instructions provide students with a rationale and, whenever possible, with an authentic task. | • Course-level assessments are in a prominent and easily identifiable location (i.e. labeled section). Grading percentages may be included in assessment descriptions, but there should be a distinct section detailing grading (see component 8). |
6. **There is evidence of plans for frequent formative assessments with immediate feedback** from a variety of sources (e.g., self, peer, instructor, computer generated, community.) These low-stakes, formative assessments allow students to “practice” before high-stakes summative assessments.

   - Examples of formative assessments might include use of clickers, informal writing assignments, group discussions or moderated discussion board, and ungraded or lightly-graded homework assignments.
   - Source of feedback may not always be evident.
   - While the syllabus might not describe all forms of formative assessments in detail, the syllabus makes clear that such activities will occur throughout the course. Evidence of formative assessment might depend on a fully articulated schedule.

7. **The assessments are adequately paced and scaffolded** (i.e., increasing in complexity) throughout the course, and at least one is scheduled early in the semester.

   - There should be evidence in the assessment descriptions or in the schedule that complex assignments build slowly over the semester or are continually re-examined with the introduction of new material.
   - Evidence of pace and scaffolding may depend on disciplinary knowledge.
   - Without a fully articulated schedule, it may not be possible to fully determine the pace and degree of scaffolding.

8. **Grading or student evaluation information is included in the syllabus but clearly separated from information about assessment of learning** (with the possible exception of the weight or percentage of the assessment in the overall course grade). Importantly, the grading scheme aligns with the learning objectives and supporting assessments.

   - The grading scheme should clearly reflect the importance of each learning objective. For example, if learning to write in the discipline is a key learning objective, writing assignments should dominate the grading scheme.

### Schedule

<table>
<thead>
<tr>
<th>Criterion</th>
<th>What the component looks like:</th>
<th>Ideas for where to look and examples of what to look for (not all need to be present):</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. <strong>Syllabus offers fully articulated and logically sequenced course schedule</strong>, listing topics/readings/questions in chronological order along with assignment due dates. Thus structured, the schedule allows for flexibility where appropriate. A schedule is necessary in order to fully evaluate the syllabus. A missing schedule may lead to low scores on components 6 and 7.</td>
<td>• The schedule is not merely a list of content topics. It contains enough information (e.g., topics, context, questions, dates) to guide students through the course. It also clearly indicates when additional information will be provided at a later date.</td>
<td></td>
</tr>
<tr>
<td>Criterion</td>
<td>What the component looks like:</td>
<td>Ideas for where to look and examples of what to look for (not all need to be present):</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Overall Learning Environment: Promise, Tone, Inclusivity | **The learning environment is supportive and invites students to engage in and take ownership of their own learning.** | - The positive, respectful, inviting tone is conveyed throughout the document.  
- Personal pronouns (e.g., you, we, us) are used, rather than “the students,” “the course,” or “they.”  
- The focus of the document is on learning and possibilities and not policies and punishments.  
- The syllabus contains a “promise” that will be fulfilled through mutual effort by instructor and students if the learning goals and objectives are met. Evidence for “promise” could include the following: language that emphasizes collaborative spirit; verbs that focus on what students and instructors do, not what the course, or some other abstract entity, does; clear statement of connections between course content and paths to answering “big questions.” |
| **10. The tone of the document is positive, respectful, inviting,** and directly addresses the student as a competent, engaged learner. | | - The positive, respectful, inviting tone is conveyed throughout the document.  
- Personal pronouns (e.g., you, we, us) are used, rather than “the students,” “the course,” or “they.”  
- The focus of the document is on learning and possibilities and not policies and punishments.  
- The syllabus contains a “promise” that will be fulfilled through mutual effort by instructor and students if the learning goals and objectives are met. Evidence for “promise” could include the following: language that emphasizes collaborative spirit; verbs that focus on what students and instructors do, not what the course, or some other abstract entity, does; clear statement of connections between course content and paths to answering “big questions.” |
| **11. The syllabus signposts a learning environment that fosters positive motivation,** one that promotes a learning orientation rather than a performance one. The document describes the potential value of the course in the learner’s current and post-course life (cognitive, personal, social, civic, and/or professional) in a clear and dynamic way. It clearly communicates that content is used primarily as a vehicle for learning, to understand core principles in the discipline and promote critical thinking and other significant learning objectives. | | - The course description makes clear that students will have opportunities to wonder and connect it in meaningful ways to things potentially important to them. The instructor encourages students to “discover” value in the course by giving them choices along the way, such as choices in project topics, reading assignments, grading schemes.  
- Various course components—description, objectives, schedule—frame the content through compelling, beautiful questions or big ideas.  
- The instructor uses information from pre-course questionnaires, background checks, pre-course exams, etc., to tailor the learning environment. In other words, he/she considers students’ backgrounds in designing course activities and assignments and takes steps to reach out to students who might struggle in class.  
- The student is left in control of his/her learning. For example, mastery-based grading mechanisms (e.g., criterion-referenced, task-based, and absolute grading schemes) are used rather than performance ones (e.g., grading curves and other relative or group-referenced grading schemes).  
- The instructor also provides resources or instruction related to becoming a lifelong learner, either in general or in ways specific to the discipline.  
- The syllabus de-emphasizes course policies by positioning them late in the syllabus or in a separate document and connecting them to |
<table>
<thead>
<tr>
<th>Validity</th>
</tr>
</thead>
</table>

To accurately score syllabi, the rubric assumes raters have three pieces of prior knowledge: 1) fluency with Fink’s Taxonomy of Significant Learning (Fink, 2013a), 2) clarity on our definitions of learning goals and objectives, and 3) familiarity with alignment as a course design construct.

**Fink’s Taxonomy of Significant Learning.** Our rubric is based on Fink’s notion that for significant learning to occur learners need to be engaged on multiple dimensions, including the cognitive, affective, and self-directed learning domains (2013a). Repackaging and expanding on Bloom’s Taxonomy of Educational Objectives (Anderson, et al., 2001; Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956), Fink’s Taxonomy of Significant Learning identifies “six kinds of significant learning” (Foundational Knowledge, Application, Integration, Human Dimension, Caring, and Learning How to Learn) and organizes them in a non-hierarchical fashion. Each of these six broad categories encompasses many types of modes of thought, skills, and intellectual and affective processes that reinforce each other and contribute to learning. Syllabi that score highly on our rubric attend to all six kinds of significant learning in Fink’s taxonomy.

**Definitions of Learning Goals and Objectives.** Accurate scoring also depends on understanding the distinction we make between learning “goals” and “objectives.” Though in common parlance, goals and objectives may be used interchangeably as synonyms for desired outcomes, in course design settings it is useful to distinguish between a course’s longer-ranging, but perhaps less tangible goals, and the shorter-term, measureable objectives of a course.

Learning goals are high-level and long-term. Educational developers can provoke instructors to envision goals for their students by asking them what they hope students will remember or be able to do three or
more years after taking the course (Fink, 2013). Setting goals gives instructors the opportunity to think about how their courses contribute to students’ learning as a whole, whether that be developing skills that will be used in subsequent courses or careers, introducing or honing discipline-specific modes of thinking, or inspiring student interest in a field that is new to them. Because they are long-ranging and more holistic in nature, goals are frequently articulated in aspirational and inspirational language in course descriptions or sections that communicate the course’s long-term value for the learner.

By contrast, a syllabus’s objectives are always concrete and measurable. They translate high-level goals into measurable course-level outcomes in such a way that students understand what the course intends for them to achieve. Students practice course objectives through a variety of formative, scaffolded learning activities. Course assessments then measure student mastery of those objectives.

Alignment. The last of our assumptions is that rubric users are familiar with alignment as a conceptual tool in the course development process. In a well aligned course, assessments and learning activities are directly derived from learning goals and objectives (Biggs & Tang, 2007; Fink, 2013; Wiggins & McTighe, 2005; Wulff & Jacobson 2005). Following the process of backward design, instructors begin by articulating objectives for student learning and then create learning activities (for both in and out of class time) and assessments that support the development of specific skills, attitudes, knowledge and values articulated in the objectives. Another way to think of alignment is for instructors to ask themselves if their in-class activities and homework assignments allow students to practice the knowledge and skills they value and if their assessments actually measure mastery of those skills.

Alignment is absolutely necessary in any learning-focused syllabus, but it is insufficient on its own to create a learning-focused document. For example, a math instructor who wishes students to memorize formulae could design a perfectly aligned course if the assessments measured memorization with multiple choice exams. The course would be aligned, but it wouldn’t score well according to our rubric because it fails to address multiple dimensions of significant learning across a wide range of learning objectives. The importance of designing assessments and learning activities with an eye toward alignment, and the insufficiency of alignment as a principle on its own, explains why alignment features prominently in several components of the syllabus assessment rubric, but does not count as its own criterion.

It is worth noting that even with a lack of fluency in Fink’s taxonomy or a lack of understanding of the distinction between goals and objectives or the nuances of alignment we have found that scoring tends to be quite consistent, across all components and among multiple raters, for syllabi that fall on the content end of the spectrum. But, this is not true of learning-focused syllabi, where scores vary wildly when raters do not have a clear understanding of our underlying conception of learner-centeredness.

Scoring

A sample scoring sheet is shown below. Each essential component (gold) is awarded three points, important components (silver) two, and less-important components (white) one, regardless of the strength of evidence. In other words, a rater would place a 3 in the appropriate strength-of-evidence column for component #1 and a 2 in the appropriate column for component #10. After scoring all components, each column is summed and scaled by the appropriate factor: the strong evidence sub-total is multiplied by 2, the moderate evidence sub-total is multiplied by 1, and the low evidence sub-total is multiplied by 0. This multi-directional weighting scheme is used in order to ensure that the final score reflects the presence and quality of essential components. A syllabus will not score high if, for example, it does not include meaningful objectives or fails to align the objectives with
the assessments. It could, however, score high if it exhibited strong evidence for most of the essential and important components but lacked evidence for the less-important ones, such as regular formative feedback opportunities or organization.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Component</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Goals &amp; Objectives</td>
<td>1. Learning goals encompass full range of Fink’s dimensions of significant learning</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>2. Course level learning objectives are clearly articulated and use specific action verbs</td>
<td>Strong</td>
</tr>
<tr>
<td></td>
<td>3. Learning objectives are appropriately pitched¹</td>
<td>Low</td>
</tr>
<tr>
<td>Assessment Activities</td>
<td>4. Objectives and assessments are aligned</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>5. Major summative assessment activities are clearly defined</td>
<td>Strong</td>
</tr>
<tr>
<td></td>
<td>6. Plans for frequent formative assessment with immediate feedback</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>7. Assessments are adequately paced and scaffolded</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>8. Grading information is included but separate from assessment; it is aligned with objectives</td>
<td>Low</td>
</tr>
<tr>
<td>Schedule</td>
<td>9. Course schedule is fully articulated and logically sequenced</td>
<td>Strong</td>
</tr>
<tr>
<td>Classroom Environment</td>
<td>10. Tone is positive, respectful, inviting</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>11. Fosters positive motivation, describes value of course, promotes content as a vehicle for learning</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>12. Communicates high expectations, projects confidence of success</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>13. Syllabus is well organized, easy to navigate, requires interaction</td>
<td>Low</td>
</tr>
</tbody>
</table>

**Subtotals:**

- Strong: 6x2 = 12
- Moderate: 14x1 = 14
- Low: 3x0 = 0

**Total:** 26/46

¹ This component, though important, goes unscored in the rubric, in recognition of the fact that the correct pitching of learning objectives would be difficult to assess without intimate knowledge of the discipline and curriculum.
Supplemental Rubric

We have also developed a supplemental rubric to assess the quality of learning activities. Because day-to-day classroom activities are often not evident in syllabi, we have chosen to parse this criterion out of the main rubric and leave it to the discretion of the rater—whether faculty developer or instructor—to decide if and when to apply these components.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>What the component looks like:</th>
<th>Ideas for where to look and examples of what to look for (not all need to be present):</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Activities</td>
<td>The “classroom” is a dynamic place and takes advantage of evidence-based practices.</td>
<td></td>
</tr>
<tr>
<td>14. It is clear that classroom activities, assessments, and learning objectives are aligned. In other words, the classroom learning activities directly support the assessments and help prepare students for them.</td>
<td>Not all classroom activities may be evident in the syllabus but there is some indication of the day-to-day structure of the learning environments. Red flags might include: exclusive use of a traditional lecture format when critical thinking is an objective; little reflective writing when self-discovery is an objective; canned homework assignments or multiple-choice tests when problem solving is an objective.</td>
<td></td>
</tr>
<tr>
<td>15. The learning activities are derived from evidence-based practices.</td>
<td>The instructor relies on pedagogical strategies and classroom activities that have some basis in the literature to support their efficacy.</td>
<td></td>
</tr>
<tr>
<td>16. The learning activities are likely to actively engage students in a variety of ways.</td>
<td>Students have opportunities, for example, to discuss course material, work individually and in groups, teach each other, solve problems, debate concepts, simulate scenarios, and/or reflect—individually and collectively—about the meaning of their learning experiences. Individual class periods involve multiple modes of instruction and varied activities.</td>
<td></td>
</tr>
</tbody>
</table>
When using the Supplemental Rubric, Component #14 is designated essential, #15 important, and #16 less-important; each is scored similarly to the scheme used for the main rubric. When applied, the maximum total score possible is 58 (46 for the main rubric and 12 for the supplemental one). In this scenario, content-focused syllabi might fall in the range 0–18, transitional 19–40, and learning-focused 41–58.

<table>
<thead>
<tr>
<th>Learning Activities</th>
<th>Score</th>
<th>Score</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>14. Classroom activities, assessments, and objectives are aligned</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Learning activities are derived from evidence-based practices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Learning activities likely to actively engage students</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subtotal: $x_2 = \quad x_1 = \quad x_0 = 0$

Total: /12

References


## Appendix A: Verbs for Significant Learning

<table>
<thead>
<tr>
<th>DIMENSION</th>
<th>ACTION VERBS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FOUNDATION KNOWLEDGE</strong>—WHAT KEY INFORMATION, IDEAS, PERSPECTIVES ARE IMPORTANT FOR LEARNERS TO KNOW?</td>
<td></td>
</tr>
<tr>
<td>Understanding and Remembering</td>
<td>Associate, Describe, Illustrate, Paraphrase, Repeat</td>
</tr>
<tr>
<td>— developing a full understanding of concepts to a degree that allows explanations, predictions, etc.</td>
<td>Compare, Explain, Indicate, Predict, Restate</td>
</tr>
<tr>
<td>Contrast</td>
<td>Give example, List, Recite, Tell</td>
</tr>
<tr>
<td>Define</td>
<td>Identify, Name, Recognize</td>
</tr>
<tr>
<td><strong>APPLICATION</strong>—WHAT KINDS OF THINKING, COMPLEX PROJECTS AND SKILLS ARE IMPORTANT FOR LEARNERS TO BE ABLE TO DO/MANAGE?</td>
<td></td>
</tr>
<tr>
<td>Critical Thinking</td>
<td>Analyze, Compare, Diagram, Hypothesize, Organize</td>
</tr>
<tr>
<td>— analyzing and critiquing issues and situations</td>
<td>Assess, Contrast, Differentiate, Infer, Query</td>
</tr>
<tr>
<td>Audit</td>
<td>Decipher, Dissect, Interpret, Separate</td>
</tr>
<tr>
<td>Catalog</td>
<td>Distinguish, Label, Trace</td>
</tr>
<tr>
<td>Categorize</td>
<td>Derive, Examine, Locate</td>
</tr>
<tr>
<td>Classify</td>
<td>Determine, Formulate, Measure</td>
</tr>
<tr>
<td><strong>Practical Thinking</strong></td>
<td>Advising, Choose, Diagnose, Predict, Select</td>
</tr>
<tr>
<td>— developing problem-solving and decision-making capabilities</td>
<td>Answer, Consult, Evaluate, Prescribe, Solve</td>
</tr>
<tr>
<td>Calculate, Decide, Determine</td>
<td>Give evidence, Judge, Prove, Suggest</td>
</tr>
<tr>
<td><strong>Creative Thinking</strong></td>
<td>Advise, Choose, Diagnose, Predict, Select</td>
</tr>
<tr>
<td>— creating new ideas, products, and perspectives</td>
<td>Abstract, Construct, Devise, Fabricate, Sketch</td>
</tr>
<tr>
<td>Adapt, Convert</td>
<td>Discover, Develop, Experiment, Theorize</td>
</tr>
<tr>
<td>Amend, Create</td>
<td>Draw, Envision, Refine, Transform</td>
</tr>
<tr>
<td>Author, Design</td>
<td>Envision, Improve, Write</td>
</tr>
<tr>
<td>Compose, Develop</td>
<td>Experiment, Refine, Write</td>
</tr>
<tr>
<td>Managing Complex Projects</td>
<td>Administer, Conduct, Facilitate, Organize, Summarize</td>
</tr>
<tr>
<td>— being able to coordinate and sequence multiple tasks in a single project/case and/or multiple projects/cases</td>
<td>Assign, Coordinate, Follow up, Plan, Teach</td>
</tr>
<tr>
<td>Coach, Delegate</td>
<td>Guide, Implement, Prioritize, Time-line</td>
</tr>
<tr>
<td>Communicate, Develop</td>
<td>Manage, Strategize, Train</td>
</tr>
<tr>
<td>Complete, Evaluate</td>
<td>Manage, Supervise</td>
</tr>
<tr>
<td><strong>Performance Skills</strong></td>
<td>Conduct, Employ, Operate, Set up</td>
</tr>
<tr>
<td>— developing capabilities in carrying out psychomotor activities</td>
<td>Demonstrate, Execute, Perform, Use</td>
</tr>
<tr>
<td>Do, Exhibit</td>
<td>Produce, Use</td>
</tr>
<tr>
<td><strong>INTEGRATION</strong>—WHAT CONNECTIONS SHOULD LEARNERS BE ABLE TO RECOGNIZE AND MAKE WITHIN AND BEYOND THIS LEARNING EXPERIENCE?</td>
<td></td>
</tr>
<tr>
<td>Interdisciplinary Learning</td>
<td>Associate, Concept map, Connect, Differentiate, Relate</td>
</tr>
<tr>
<td>— connecting ideas, disciplines, perspectives, contexts</td>
<td>Combine, Contrast/compare, Correlate, Link, Synthesize</td>
</tr>
<tr>
<td>Learning Communities</td>
<td>— connecting people</td>
</tr>
<tr>
<td>Learning and Living/Working</td>
<td>— connecting different realms of life</td>
</tr>
</tbody>
</table>

**HUMAN DIMENSION—WHAT SHOULD LEARNERS LEARN ABOUT THEMSELVES AND ABOUT INTERACTING WITH OTHERS?**

<table>
<thead>
<tr>
<th>Interpersonal Relationships</th>
<th>Acquire</th>
<th>Describe</th>
<th>Inspire</th>
<th>Protect</th>
<th>Unite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Authorship</td>
<td>Advise</td>
<td>Demonstrate</td>
<td>Interact with</td>
<td>Reconcile</td>
<td>Critically</td>
</tr>
<tr>
<td>Leadership</td>
<td>Advocate</td>
<td>Educate</td>
<td>Involve</td>
<td>Reform</td>
<td>reflect</td>
</tr>
<tr>
<td>Ethics, Character Building</td>
<td>Balance</td>
<td>Embody</td>
<td>Lead</td>
<td>Respect</td>
<td>Resolve</td>
</tr>
<tr>
<td>Multicultural Education</td>
<td>Be aware of</td>
<td>Empathize</td>
<td>Mediate</td>
<td>Settle</td>
<td>conflict</td>
</tr>
<tr>
<td>Working as a Member of a Team</td>
<td>Collaborate</td>
<td>Express</td>
<td>Mobilize</td>
<td>Share</td>
<td>Respond</td>
</tr>
<tr>
<td>Citizenship</td>
<td>Comply</td>
<td>Feel confident</td>
<td>Motivate</td>
<td>Show</td>
<td>sensibly</td>
</tr>
<tr>
<td>Environmental Ethics</td>
<td>Communicate</td>
<td>Give feedback</td>
<td>Negotiate</td>
<td>Suggest</td>
<td>Serve as role model</td>
</tr>
<tr>
<td></td>
<td>Cooperate</td>
<td>Help</td>
<td>Offer</td>
<td>Support</td>
<td>Sustain</td>
</tr>
<tr>
<td></td>
<td>Decide to</td>
<td>Influence</td>
<td>Promote</td>
<td>Sustain</td>
<td>Suspend</td>
</tr>
</tbody>
</table>

**CARING—WHAT CHANGES IN LEARNERS’ FEELINGS, INTERESTS, VALUES ARE IMPORTANT?**

<table>
<thead>
<tr>
<th>Agree to</th>
<th>Be ready to</th>
<th>Be willing to</th>
<th>Be responsible to</th>
<th>Be available to</th>
<th>Be adaptable to</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wanting to Be a Good Learner</td>
<td>Commit to</td>
<td>Decide to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Becoming Excited About a Particular Activity/Subject</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Developing a Commitment to Live Right (i.e., deciding to take care of one’s health/well-being, live by a certain code)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**LEARNING HOW TO LEARN—WHAT SHOULD LEARNERS LEARN ABOUT LEARNING, ENGAGING IN INQUIRY, AND BECOMING SELF-DIRECTED?**

<table>
<thead>
<tr>
<th>How to Be a Better Learner</th>
<th>Describe how to Research</th>
<th>Construct knowledge about</th>
<th>Identify sources and resources for</th>
<th>Take responsibility for knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>How to Inquire and Construct Knowledge</td>
<td>Inquire</td>
<td>Develop a learning plan</td>
<td>Identify what you need to know</td>
<td>Transfer knowledge</td>
</tr>
<tr>
<td>How to Pursue Self-Directed or Intentional Learning</td>
<td>Reflect</td>
<td>Frame useful questions</td>
<td>Predict performance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-assess</td>
<td>Generalize knowledge</td>
<td>Set a learning agenda</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-regulate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-monitor</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B: Blank Scoring Sheet

Award each essential component (gold) three points, important components (silver) two, and less-important components (white) one, regardless of the strength of evidence. For example, raters should place a 3 in the appropriate strength-of-evidence column for component #1 and a 2 in the appropriate column for component #10. After scoring all components, sum and scale each column by the appropriate factor: multiple the strong evidence subtotal by 2, the moderate evidence subtotal by 1, and the low evidence subtotal by 0.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Component</th>
<th>Strength of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Goals &amp; Objectives</td>
<td>1. Learning goals encompass full range of Fink’s dimensions of significant learning</td>
<td>Strong  Moderate  Low</td>
</tr>
<tr>
<td></td>
<td>2. Course level learning objectives are clearly articulated and use specific action verbs</td>
<td>Strong  Moderate  Low</td>
</tr>
<tr>
<td></td>
<td>3. Learning objectives are appropriately pitched</td>
<td>Strong  Moderate  Low</td>
</tr>
<tr>
<td>Assessment Activities</td>
<td>4. Objectives and assessments are aligned</td>
<td>Strong  Moderate  Low</td>
</tr>
<tr>
<td></td>
<td>5. Major summative assessment activities are clearly defined</td>
<td>Strong  Moderate  Low</td>
</tr>
<tr>
<td></td>
<td>6. Plans for frequent formative assessment with immediate feedback</td>
<td>Strong  Moderate  Low</td>
</tr>
<tr>
<td></td>
<td>7. Assessments are adequately paced and scaffolded</td>
<td>Strong  Moderate  Low</td>
</tr>
<tr>
<td></td>
<td>8. Grading information is included but separate from assessment; it is aligned with objectives</td>
<td>Strong  Moderate  Low</td>
</tr>
<tr>
<td>Schedule</td>
<td>9. Course schedule is fully articulated and logically sequenced</td>
<td>Strong  Moderate  Low</td>
</tr>
</tbody>
</table>

Subtotals: $x_2 \times$ Moderate $x_1 \times$ Low $x_0 = 0$
TOTAL: /46

| Classroom Environment | 10. Tone is positive, respectful, inviting                                   | Strong  Moderate  Low |
|                      | 11. Fosters positive motivation, describes value of course, promotes content as a vehicle for learning | Strong  Moderate  Low |
|                      | 12. Communicates high expectations, projects confidence of success          | Strong  Moderate  Low |
|                      | 13. Syllabus is well organized, easy to navigate, requires interaction     | Strong  Moderate  Low |

Learning Activities

| 14. Classroom activities, assessments, and objectives are aligned | Strong  Moderate  Low |
| 15. Learning activities are derived from evidence-based practices | Strong  Moderate  Low |
| 16. Learning activities likely to actively engage students      | Strong  Moderate  Low |

Subtotal: $x_2 \times$ Moderate $x_1 = x_0 = 0$
TOTAL: /12

Content-focused syllabi typically fall in the range 0–16, transitional 17–30, and learning-focused 31–46 (or 0–18, 19–40, and 41–58, respectively, when using the supplemental rubric).
Appendix B
## Syllabus Evaluation Rubric

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Beginning</th>
<th>Emerging</th>
<th>Exemplary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Course Description</strong></td>
<td>Instructor name &amp; contact info, class time and location</td>
<td>in addition: course prerequisites (if any) course description</td>
<td>in addition, how the course fits into the larger program/department curriculum, field, supplemental readings, and resources</td>
</tr>
<tr>
<td><strong>Overall Tone</strong></td>
<td>Mechanical, dictatorial</td>
<td>teacher-oriented</td>
<td>student/learning oriented (eg: first person)</td>
</tr>
<tr>
<td><strong>Course Outcomes</strong></td>
<td>not articulated</td>
<td>stated in general, but vague and unmeasurable terms</td>
<td>listed with appropriate, descriptive verbs that lend themselves to measurement and seek higher levels of learning</td>
</tr>
<tr>
<td><strong>Course Format</strong></td>
<td>vague, or criptic descriptions of course expectations and how class time will be used</td>
<td>mutual role expectations for students and instructor are explained, together with various teaching methods and modes</td>
<td>role expectations and class format are explained in such a way that students understand the underlying rationale and benefits for them</td>
</tr>
<tr>
<td><strong>Instructor Beliefs &amp; Assumptions</strong></td>
<td>little or no accounting of the instructor’s teaching philosophy, beliefs or assumptions about learning</td>
<td>section describing the instructor’s beliefs or assumptions about teaching and learning that guide the course</td>
<td>well articulated and thought out rationale that includes the values and/or experiences that guide the instructor’s teaching practice</td>
</tr>
<tr>
<td><strong>Class Schedule</strong></td>
<td>little or no information on what course topics will be covered each week</td>
<td>course topics broken down by class period</td>
<td>fully articulated and logically sequenced course schedule with chronological topics listed for each class, along with required readings and preparation necessary from students</td>
</tr>
<tr>
<td><strong>Assignments Required</strong></td>
<td>course assignments listed but with no due dates</td>
<td>course assignments listed with clear due dates</td>
<td>assignments listed with due dates, with explanation of late policy and other requirements that might affect grades</td>
</tr>
</tbody>
</table>

---

**Center for Teaching Excellence**
<table>
<thead>
<tr>
<th>Academic Policies &amp; Procedures</th>
<th>little or no information</th>
<th>description of academic integrity policy</th>
<th>information about all pertinent academic policies, including academic integrity, accommodating students with disabilities, class attendance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment of Students' Learning</td>
<td>little or no information about how the students will be graded; any information that is included reinforces a grade-focus</td>
<td>Each graded assignment is clearly described with its relative value towards the overall course grade</td>
<td>Each assignment includes descriptions of its rationale for inclusion in the course and what the student should get out of completing it; use of rubrics with quality criteria specified</td>
</tr>
<tr>
<td>Alignment</td>
<td>no clear connection between stated course goals/outcomes and assessment schema</td>
<td>the connections between some assignments and stated course goals/outcomes are apparent</td>
<td>all assignments are linked with a specific course goal/outcome and are likely to provide sufficient evidence to adequately assess each goal/outcome</td>
</tr>
<tr>
<td>Diversity of Teaching &amp; Assessment Methods</td>
<td>course teaching and assessment methods are similar; eg: all lectures; all tests</td>
<td>Evidence the instructor has employed a diverse set of teaching and assessment methods</td>
<td>Diverse assessment methods and evidence that the instructor has taken into account the diversity of students in choosing teaching and assessment methods</td>
</tr>
<tr>
<td>Continuity of Feedback to Students on Their Learning</td>
<td>little or very infrequent venues for giving students feedback on their progress in the course</td>
<td>adequate opportunities for students to get feedback on their progress in the course</td>
<td>all course requirements have sufficient means by which the instructor can keep students adequately appraised of their relative progress in the course</td>
</tr>
<tr>
<td>Opportunities for Students to Provide Evaluative Course Input</td>
<td>students' only opportunity to provide input on their experiences in the course to the instructor is at the end of the course</td>
<td>Instructor has developed and scheduled a mid-semester course evaluation opportunity for the students</td>
<td>Students are encouraged to provide the instructor with regular input on how they are experiencing the course throughout the semester</td>
</tr>
</tbody>
</table>
Appendix C
WRITING A STATEMENT OF TEACHING PHILOSOPHY
FOR THE ACADEMIC JOB SEARCH

Chris O’Neal, Deborah Meizlish, and Matthew Kaplan

Domestic Environmental Policy and Politics. Lehigh University’s year-old Environmental Initiative seeks an Assistant Professor for a tenure track position... To apply, please send a cover letter, current curriculum vitae, syllabi and other evidence of teaching style and effectiveness, a statement of teaching philosophy, a sample of scholarship (if available) and three letters of reference.

Assistant Professor (tenure track) Specialization in African and Post-Colonial Literatures.... Send letter of application, curriculum vitae, statement of teaching philosophy, graduate school transcript, and three letters of recommendation... Northeastern Illinois University is an affirmative action, equal opportunity employer.

LSU’s Department of Chemistry (chemistry.lsu.edu) anticipates filling one or two tenure-track positions in the fields of NMR Spectroscopy (Ref: Log #0184) and Physical Chemistry (Ref: Log #0186), broadly defined.... Applications should consist of a research proposal, a statement of teaching philosophy, and a curriculum vitae (including address). Applicants should arrange for submission of three letters of recommendation.

Introduction

As these recent job ads illustrate, requests for teaching philosophies are common in the academic market. In fact, a survey of 457 search committee chairs in six disciplines (English, history, political science, psychology, biology, and chemistry) found that 57% requested a teaching statement at some point in a job search (Meizlish & Kaplan, in press). These results differed slightly by institutional type, with master’s and bachelor’s institutions requesting them more often than doctoral institutions. Results also differed by discipline. Surprisingly, requests for teaching philosophies were most frequent in the natural sciences. But the overall message is clear: job applicants in all fields may be asked to submit a teaching philosophy (see also Bruff, in press; Montell, 2003; Schönwetter, Taylor, & Ellis, 2006).

Chris O’Neal is Senior Consultant for Institutional Initiatives at the Center for Research on Learning and Teaching (CRLT). Deborah Meizlish is Coordinator of Social Science Initiatives at CRLT. Matthew Kaplan is Managing Director of CRLT. They have Ph.D.s in Biology, Political Science, and Comparative Literature, respectively.
Teaching philosophies can serve several purposes (e.g., self-reflection, introduction to a teaching portfolio, communication with students), but we focus here on those written for academic job applications. Such statements communicate a job candidate’s approach to teaching and learning to a faculty considering whether to make that candidate one of their colleagues. Since a committee cannot possibly observe the teaching of every applicant, the teaching philosophy helps search committee members imagine themselves in each candidate’s classroom. What is it like to be one of this instructor’s students? Why does she make the pedagogical decisions she does? As a student in this classroom, how would I spend my fifty minutes on a given day? How does the instructor address the challenges and resources of teaching in his particular discipline? Does her teaching style complement our department’s philosophy of instruction?

This Occasional Paper is designed to help experienced graduate students write a statement of teaching philosophy. The paper contains four sections. First, we offer suggestions for making a philosophy of teaching explicit and getting it on paper. Second, we discuss research on characteristics of effective statements. Third, we introduce a rubric that can guide the development and crafting of a teaching statement that search committees will value. Finally, we address questions that job candidates often raise about this sometimes perplexing document.

Advice for Getting Started

Just because you have never written a statement of your teaching philosophy does not mean you do not have a philosophy. If you engage a group of learners who are your responsibility, then your behavior in designing their learning environment must follow from your philosophical orientation. What you need to do is discover what [your philosophy] is and then make it explicit. (Coppola, 2000, p. 1)

Beginning the teaching philosophy is often the hardest part of writing one. The motivations behind the decisions we make in the classroom can be surprisingly elusive when we try to put them on paper. Since there is no single approach that will work for all writers, we offer three strategies for getting started:

1. Goodyear and Allchin (1998) found that thinking about the “big” questions of teaching helped instructors articulate their philosophies:
   - What motivates me to learn about this subject?
   - What do I expect to be the outcomes of my teaching?
   - How do I know when I’ve taught successfully?

2. In workshops and seminars at U-M, we have found that some graduate students prefer to approach a statement by thinking about more concrete and manageable "fragments" of teaching that can then be assembled into a holistic essay. The following questions are designed to get you started:
   - Why do you teach?
   - What do you believe or value about teaching and student learning?
   - If you had to choose a metaphor for teaching/learning, what would it be?
   - How do your research and disciplinary context influence your teaching?
   - How do your identity/background and your students’ identities/backgrounds affect teaching and learning in your classes?
   - How do you utilize multiple pedagogical approaches in your teaching?
   - What is your approach to evaluating and assessing students?

3. Finally, some instructors find it most useful to begin by simply looking at examples of others’ philosophies. CRLT has posted sample statements from a variety of disciplines at <http://www.crlt.umich.edu/tstrategies/tstpum.html>. When looking at others’ philosophies, you will likely note considerable variation, both in terms of content and format, and you will likely find some approaches that resonate with you. While there is no single approach to a teaching philosophy, Figure 1 provides some general guidelines for those statements written for the academic job market.

Figure 1. Some general guidelines for writing the teaching philosophy (adapted from Chism, 1998):

- Keep it brief (1–2 pages).
- Use a narrative, first person approach.
- Make it reflective and personal.
- Discuss your goals for your students, the methods you use to achieve those goals, and the assessments you use to find out if students have met your expectations.
- Explain your specific disciplinary context and use specific examples of your practice.
- Showcase your strengths and accomplishments.
Once you’ve articulated a first draft, you can begin shaping and polishing it for the search committees who will be reading it. In the following section, we discuss characteristics of successful teaching philosophy statements and provide a rubric for evaluating a teaching statement and aiming it at the right audience.

**What Constitutes a Good Statement?**

In their survey of search committee chairs, Meizlish and Kaplan (in press) found broad agreement on the desirable characteristics of a statement of teaching philosophy. Specifically, chairs described successful teaching statements as having the following characteristics:

- **They offer evidence of practice.** Search committee chairs want to understand how candidates enact their teaching philosophies. In particular, they want to see specific and personal examples and experiences rather than vague references to educational jargon or formulaic statements.

- **They convey reflectiveness.** Search committees want to know that a candidate is a thoughtful instructor. They are interested in candidates who can discuss their approach to instructional challenges and their plans for future pedagogical development.

- **They communicate that teaching is valued.** Search chairs appreciate a tone or language that conveys a candidate’s enthusiasm and commitment to teaching. They are wary of candidates who talk about teaching as a burden or a requirement that is less important than research.

- **They are student- or learning-centered, attuned to differences in student abilities, background knowledge, or levels.** Search committee chairs want concrete evidence of a candidate’s attentiveness to student learning (rather than just content) and awareness of and ability to deal with student differences in the classroom.

- **They are well written, clear, and readable.** Search chairs draw conclusions about candidates from all elements of the application packet. Candidates can be undermined by carelessness in their teaching statements.

**A Rubric to Evaluate the Teaching Philosophy**

Based on survey responses from search committee chairs, our own experience reading hundreds of teaching philosophies, and research on best practices in teaching and learning, we constructed a rubric to help graduate students write and evaluate statements of teaching philosophy (Figure 2). The rubric can be used as a starting point for revising first drafts of your philosophy. The rubric consists of the following five categories:

1. Goals for student learning
2. Enactment of goals
3. Assessment of goals
4. Creating an inclusive learning environment
5. Structure, rhetoric, and language

The first three categories of the rubric were purposefully framed to encourage instructors to think about the alignment of their goals, methods, and assessments. Research suggests that aligning intended outcomes (goals), instructional methods, and testing can lead to significant gains in student learning. Instructional alignment is more important for tasks involving higher-order thinking skills, and it has a particularly strong impact on the performance of lower aptitude students (Cohen, 1987).

In terms of writing a teaching statement, focusing on alignment raises a number of useful questions about your approach to teaching and student learning: What do you want students to learn (and why)? What approach will you take to help students acquire the desired knowledge, skills, and attitudes, and how can you best test students to determine whether they have reached these goals? Reflecting on these issues in a systematic fashion allows you to develop a clear sense of why you take the approach you do, often one of the most difficult aspects of writing a teaching statement. It also has the potential to reveal areas of misalignment, providing clear direction for future development and ensuring that the teaching statement is not merely a rhetorical exercise, but a useful contribution to your development as a teacher.

The fourth category reflects our belief that pedagogical practices that reach students at the margins of the classroom are beneficial for all students. As Kardia (1998) writes,

- **Attention to race/ethnicity, gender, sexual orientation, disability, and other student characteristics is consistent**
with an improved learning environment for all students. For example, an instructor who provides more lead time with assignments in response to the needs of a student with a physical disability will be appreciated by all students, even though the majority of the students might have been able to find ways to compensate for the lack of lead time. (p. 19)

Research has confirmed the benefits of diversity for promoting student learning and development. For example, studies conducted at U-M on the impact of racial diversity on student learning and attitudes confirm that positive classroom interactions across racial difference can lead to increased student motivation, critical thinking skills, and social engagement. Obviously, it is up to faculty to create positive learning experiences in order to take advantage of diversity. "Students, indeed, acquire a very broad range of skills, motivations, values, and cognitive capacities from diverse peers when provided with the appropriate opportunities to do so" (Gurin, 1999, Conclusion). Future faculty need to demonstrate that they have thought carefully about these issues.

The last category addresses some of the most common complaints search committee chairs voiced about teaching statements. Chairs complained about teaching jargon that alienates many readers and weak thematic structures that make reading difficult. Obviously, search committees are more likely to have a positive view of a well-written teaching philosophy than a poorly written one.

A common component running through all of these categories is a focus on specificity, disciplinary context, and rich, illustrative examples. The importance of this component is based on the finding that search committees want to read about specific examples of how candidates enact their teaching philosophies. What does this mean in practice? Rather than saying, “I use active learning in my teaching,” write about a specific exercise you use in your class that engages students actively. Why do you use it? How were students different after the activity? Thinking about your students, what do they typically find most challenging about that activity? How do you know that the activity worked?

Below we provide excerpts from teaching philosophies written by U-M graduate students that exemplify each of the rubric’s first four categories (the fifth, by necessity, is illustrated by each example).

**Goals for student learning**

At the heart of most teaching philosophies is a set of goals for what an instructor hopes to accomplish in the classroom. An instructor’s goals should describe how students will be different after leaving that instructor’s class. What will they be capable of doing that they could not before? What will they know that they did not before? How will they see the world differently? Goals in a teaching philosophy should be clearly written to describe the ways students will develop, as well as to convey the context of the instructor’s discipline.

In this description of goals, a social work graduate student instructor (GSI) talks about the transformative nature of social work education. Her goals for her students are lifelong and directly tied to the mission of social work as a discipline. Note the specificity of the skills she hopes students will attain.

Social work education should foster students’ critical consciousness – the ability "to perceive social, political, and economic contradictions and to take action against those oppressive elements" (Freire).... Through creative and interactive activities in and outside of my classroom, students learn to recognize, analyze, and work to change dynamics of privilege and oppression when engaging with others in all areas of practice – individual, group, community, or state.

**Enactment and assessment of goals**

A teaching philosophy cannot rest solely on an instructor's learning goals. For it to be useful in the job search, it must also communicate how instructors will achieve those goals, and how they will know that they have achieved them. Specificity is compelling when talking about teaching methods and assessments. Likewise, the more closely the methods and assessments are grounded in disciplinary pedagogies, the more they will resonate with readers in that field. The first example below describes the teaching methods used by a GSI in Germanic Languages and Literatures. The second describes how an electrical engineering GSI assesses student learning.

While confronting my students with the challenge of learning new languages and cultures, I encourage them to reflect on their own beliefs and try to open their minds to new ways of seeing things. For example, in my
fourth-semester German class that focused on the lives of and work of the Brothers Grimm, we often engaged in discussion of original, European fairy tales, comparing them to their well-known Disney versions. Such comparisons helped my students not only to learn about important aspects of German literature, but also encouraged them to step back and reflect on the values of their own culture.

In order to solve new problems, engineers should be able to think through them. The final solution to a problem is rarely obvious and, as such, the thinking process must be developed and refined with practice. In a term, I assign several individual and group projects that incorporate multiple ideas and first principles. Projects early in a term are broken down with milestones such that students can begin to learn how to approach a multifaceted problem on their own…. In addition to projects, students use their critical thinking skills on a more regular basis during weekly timed quizzes. The quizzes are not designed to test memorization…rather, they are designed to test problem solving, as each quiz cannot be completed if not approached properly.

Creating an inclusive learning environment

This category emphasizes the integration of inclusive teaching and learning throughout the statement, thereby avoiding the isolated “diversity paragraph.” In the following quotations, the authors connect inclusive teaching to their goals for their courses and their understandings of their disciplines.

Parallel to the idea of discovering new things as an engineer is the idea of discovering new minds and cultures. Similarly, learning analytical and evaluation skills as an engineer parallels learning to understand and/or tolerate other points of view…. In my classes I try to expose the students to different situations to help them gain these skills, including interacting with classmates with different backgrounds (race, ethnicity, gender, technical knowledge...), taking different roles when working in teams (leader, note taker, report writer, etc.), and taking different roles when working individually (presenter or evaluator). By doing so, I hope to provide the students the opportunity to learn not only the theory of mechanical engineering and problem solving skills, but also to realize that around them there is much to learn as well. (GSI in Mechanical Engineering)

My ideal classroom is primarily a safe and comfortable place where students of diverse background and experience are encouraged to clarify their thoughts and expose their assumptions…for mutual examination…. On whichever level I teach, I intend that my courses are enriching to my students of diverse background in various ways that will suit their particular academic and personal needs. Specifically, through the activities discussed above, I hope that those students of general North American cultural background broaden their intellectual and spiritual horizons by critically reflecting upon their own cultural assumptions and beliefs and the students of Asian ancestry who are curious about their own philosophical and religious traditions can enrich themselves by learning more about their roots. (GSI in Asian Languages and Culture)

A statement need not achieve a rating of “excellent” in each of the categories described in the rubric to be a good teaching statement. We encourage you to seek input on your teaching statement in much the same way you would solicit feedback on a scholarly paper. Faculty in your department can provide feedback based on their own experience serving on search committees and reading application materials. Keep in mind, however, that the qualities that serve a job candidate well at U-M may not match those at a different institution. For a different perspective, you might ask for feedback from peers in your program who have graduated and are now faculty elsewhere or from mentors at your undergraduate institution.

Frequently Asked Questions

1. Do I have to write a new teaching philosophy for every school?

An individual's core teaching philosophy probably will not change based on the school to which he or she is applying for a position. That said, search committees are attentive to the match or mismatch between the priorities of their institution or department and the priorities implied by a job candidate's teaching statement and other application materials. It is worth considering the range of positions to
which you are applying and thinking carefully about whether some degree of customization (based on institutional type, focus of the position, etc.) is appropriate.

2. What should I do if I don’t have a lot of teaching experience upon which to base my statement of teaching philosophy?

This is not an uncommon situation, particularly in some disciplines where teaching opportunities for graduates are rare. Regardless of your experience as an instructor, you have years of experience as a student in your discipline that you can draw from. Additionally, you may have some experience mentoring students in the lab, independent study, or elsewhere. Talk about your approach to teaching in these settings and the lessons you would take to your own classroom.

3. Can sending an unsolicited teaching philosophy hurt me when I’m applying for faculty positions?

Meizlish and Kaplan asked search committees this very question. The conclusion was clear: submitting an unsolicited teaching statement is viewed positively by most search committee chairs.

4. Should I reference or include student ratings and comments?

A teaching statement is a brief overview of your approach to teaching supported by rich examples drawn from your practice. As a result, support materials such as student ratings and comments would be out of place in a teaching statement. Consider instead constructing a teaching portfolio to highlight these materials. Teaching portfolios are organized, annotated collections of the “evidence” that supports your philosophy. They can include student evaluations, samples of assignments, letters of recommendation, samples of student work, etc. Note that teaching portfolios are rarely requested by search committees. For more information, see CRLT Occasional Paper No. 11, The Teaching Portfolio (available at http://www.crlt.umich.edu/resources/occasional).

5. Are teaching philosophies original work? Couldn’t I adopt someone else’s philosophy if I completely agree with what they’re saying?

Teaching philosophies are original work, just like anything else you or someone else writes. Copying others’ philosophies is plagiarism. Besides, a well-written philosophy should be rooted in your own practice and illuminated by specific examples from your own work. No one else has had exactly your experiences in the classroom.

6. Will this be the last time I write a teaching philosophy?

Teaching philosophies are becoming a common component of tenure and promotion packages at colleges and universities. If you continue in academia as a tenured or untenured faculty member, a teaching statement will likely be one of the ways in which your performance is assessed. Fortunately, having written one for the job search, you will have a head start. Remember, however, that the teaching philosophy is an evolving document, changing as you gain more experience as a teacher and your beliefs about effective teaching and learning evolve. Returning to the teaching philosophy statement throughout your career is a useful reflective exercise that can help to make your current teaching practice more explicit and deliberate.

7. Where can I learn more about teaching philosophies?

The CRLT Teaching Strategies website contains a section on teaching statements (http://www.crlt.umich.edu/tstrategies/tstpts.html) with useful articles and sample statements from a variety of disciplines. CRLT offers workshops on writing teaching statements at a one-day Preparing Future Faculty Conference each fall. Graduate students interested in a more intensive experience can apply to participate in a month-long Preparing Future Faculty Seminar that is co-sponsored by Rackham and offered every May. (See http://www.crlt.umich.edu/gsis/pff.html for more information about these programs.) CRLT’s Graduate Teaching Consultants are also available to consult one-on-one about teaching philosophies. You can contact CRLT (764-0505, crlt@umich.edu) to set up a consultation.
Figure 2. Rubric for composing and evaluating a statement of teaching philosophy

<table>
<thead>
<tr>
<th>Categories</th>
<th>Excellent</th>
<th>Needs Some Revision</th>
<th>Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Goals for student learning:</strong></td>
<td>Goals are clearly articulated, specific, and go beyond knowledge level, including skills, attitudes, career goals, etc. Goals are sensitive to the context of the instructor’s discipline. They are concise, not exhaustive.</td>
<td>Goals are articulated but may be too broad or not specific to the discipline. Goals focus on basic knowledge, ignoring skills acquisition and affective change.</td>
<td>Articulation of goals is unfocused, incomplete, or missing.</td>
</tr>
<tr>
<td>What knowledge, skills, and attitudes are important for student success in your discipline?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What are you preparing students for? What are key challenges in the teaching-learning process?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Enactment of goals (teaching methods):</strong></td>
<td>Enactment of goals is specific and thoughtful. Includes details and rationale for teaching methods. The methods are clearly connected to specific goals and are appropriate for those goals. Specific examples of the methods in use within the disciplinary context are given.</td>
<td>Description of teaching methods not clearly connected to goals, or if connected, not well developed (seems like a list of what is done in the classroom). Methods are described, but generically; no example of the instructor’s use of the methods within the discipline is communicated.</td>
<td>Enactment of goals is not articulated. If there is an attempt at articulating teaching methods, it is basic and unreflective.</td>
</tr>
<tr>
<td>What teaching methods do you use? How do these methods contribute to your goals for students? Why are these methods appropriate for use in your discipline?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Assessment of goals (measuring student learning):</strong></td>
<td>Specific examples of assessment tools are clearly described. Assessment tools are aligned with teaching goals and teaching methods. Assessments reinforce the priorities and context of the discipline both in content and type.</td>
<td>Assessments are described, but not connected to goals and teaching methods. Description is too general, with no reference to the motivation behind the assessments. There is no clear connection between the assessments and the priorities of the discipline.</td>
<td>Assessment of goals is not articulated or mentioned only in passing.</td>
</tr>
<tr>
<td>How do you know your goals for students are being met? What sorts of assessment tools do you use (e.g., tests, papers, portfolios, journals), and why? How do assessments contribute to student learning? How do assessments communicate disciplinary priorities?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Creating an inclusive learning environment, addressing one or more of the following questions:</strong></td>
<td>Portrays a coherent philosophy of inclusive education that is integrated throughout the statement. Makes space for diverse ways of knowing and/or teaching approaches. Discussion of roles is sensitive to historically underrepresented students. Demonstrates awareness of issues of equity within the discipline.</td>
<td>Inclusive teaching is addressed but in a cursory manner or in a way that isolates it from the rest of the philosophy. Author briefly connects identity issues to aspects of his/her teaching.</td>
<td>Issues of inclusion are not addressed or addressed in an awkward manner. There is no connection to teaching practices.</td>
</tr>
<tr>
<td>How do your own and your students’ identities (e.g., race, gender, class), backgrounds, experiences, and levels of privilege affect the classroom? How do you use multiple teaching approaches? How do you integrate diverse perspectives into your teaching?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Structure, rhetoric and language:</strong></td>
<td>The statement has a guiding structure and/or theme that engages the reader and organizes the goals, methods, and assessments articulated in the statement. Jargon is avoided and teaching terms (e.g., critical thinking) are given specific definitions that apply to the instructor’s disciplinary context. Grammar and spelling are correct.</td>
<td>The statement has a structure and/or theme that is not connected to the ideas actually discussed in the statement, or, organizing structure is weak and does not resonate within the disciplinary context. The statement contains some jargon.</td>
<td>No overall structure present. Statement is a collection of disconnected statements about teaching. Jargon is used liberally and not supported by specific definitions or examples. Needs much revision.</td>
</tr>
<tr>
<td>How is the reader engaged? Is the language used appropriate to the discipline? How is the statement thematically structured?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Center for Research on Learning and Teaching (CRLT), University of Michigan
References


The CRLT Occasional Papers series is published on a variable schedule by the Center for Research on Learning and Teaching at the University of Michigan. Information about extra copies or back issues can be obtained by writing to: Publications, CRLT, 1071 Palmer Commons, 100 Washtenaw Avenue, Ann Arbor, MI 48109-2218.

Copyright 2007 The University of Michigan

CRLT Occasional Paper No. 23

Center for Research on Learning and Teaching

The University of Michigan
1071 Palmer Commons
100 Washtenaw Avenue
Ann Arbor, MI 48109-2218
THE TEACHING PORTFOLIO

Matthew Kaplan

At institutions across the country, faculty are creating opportunities to exchange ideas on teaching and, in the process, becoming more reflective about their teaching. In part, this is a response to national discussions about the false dichotomy that is often drawn between teaching and research. To move beyond this debate, there have been calls for expanding the idea of scholarship to include certain teaching products, as well as research products (Boyer, 1990). Three strategies for taking a scholarly approach to reviews of teaching are ones that are common to discussions of research as well (Shulman, 1993). First, scholarship is firmly grounded in the disciplines, and a scholarly approach to the review of teaching would focus on the teaching of a specific discipline. Second, just as research becomes scholarship when it is shared, faculty would need to begin making teaching community property. And finally, scholarship often involves making judgments about faculty work, which, for teaching, would mean that faculty would become more involved in reviewing each others’ accomplishments in teaching and learning.

The teaching portfolio is one of the tools faculty can use to document their scholarly work in teaching. This Occasional Paper contains a discussion of the nature and purpose of the teaching portfolio (and its offshoot, the course portfolio) and suggestions for how individuals and units can use portfolios most effectively.

What Is a Teaching Portfolio?

A record of accomplishments in teaching

Based on the model of the portfolio kept by artists and architects, the teaching portfolio contains evidence of a faculty member’s achievements in teaching: “What is a teaching portfolio? It includes documents and materials which collectively suggest the scope and quality of a professor’s teaching performance...The portfolio is not an exhaustive compilation of all of the documents and materials that bear on teaching performance. Instead, it presents selected information on teaching activities and solid evidence of their effectiveness” (Seldin, 1997, p. 2).

Documentation in context

The portfolio should be more than a simple collection of documents.

Matthew Kaplan is an instructional consultant in the Center for Research on Learning and Teaching.
It also should contain reflective statements on the material included and on the faculty member's approach to teaching and student learning. The reflective portions of the portfolio help set the documents in context for the reader; the materials provide evidence to back up the assertions made in the reflective statement.

What Might Go into a Portfolio?

When considering the contents of a portfolio, faculty must distinguish clearly between being representative and being exhaustive. Attempts to create an exhaustive compendium of an instructor's work in teaching run the risk of becoming exhausting, both for the person collecting the materials and for any readers who might choose (or need) to respond to the portfolio. Furthermore, the attempt to be completely comprehensive can turn the project of developing a portfolio into a paper chase. Such a large collection of documents makes it difficult to maintain the reflective aspect of the portfolio, which is one of its chief purposes and advantages.

The portfolio should, instead, be representative of the various aspects of a faculty member's teaching. This means looking beyond the most obvious part of teaching — what goes on in the classroom. While the activities and interactions with students in class are important, they do not fully reflect faculty work with teaching. Other items might include planning courses, assessing student learning, advising students (in office hours or in larger projects such as theses and dissertations), curriculum development and assessment, supervising student research, working to improve one's teaching, and publishing articles on teaching and learning.

One way to categorize items that a faculty member might include is to divide them into three categories based on the source of the item: materials from oneself (e.g., reflective statements, descriptions of course responsibilities, syllabi, assignments), materials from others (e.g., statements from colleagues who have observed or reviewed teaching materials, student ratings, letters from students or alumni, honors or recognition); and products of good teaching (student essays or creative work, a record of students who have succeeded in the field, evidence of supervision of theses). Some of these sources may be more appropriate for certain aspects of teaching than for others. See Appendix A for a more comprehensive list.

**Purposes of Portfolios**

**Self-reflection and improvement**

Assembling a portfolio involves reflection. Most portfolios include a reflective statement that can cover topics such as the instructor's approach to teaching and learning, his or her assumptions about the roles of students and teachers, and goals the instructor expects students to achieve (Chism, 1997-1998). In addition, faculty need to collect documents that support their reflective statement, a process that also involves reflection (selecting some items over others, reviewing past work, etc.). As a result, the portfolio is well-suited to helping faculty examine their goals for teaching and student learning, and compare those goals to the reality of their praxis.

The comparison between the ideal and the real is the first step in the process of improving teaching. Instructors can gain a sense of how effective their teaching is and how they could improve from a variety of sources: student ratings of instruction, midsemester feedback, self-perception, discussions with colleagues, etc. By constructing a portfolio, faculty will look systematically at the various sources of data about their teaching; therefore, they can make more informed decisions about teaching strengths on which they wish to build and problems in their teaching they wish to address. The reflection and improvement process can be further enhanced when faculty work together (in pairs or small groups) as they develop their portfolios. Colleagues can offer support and advice, exchange new ideas and solutions to problems, and broaden each other's views of the teaching and learning process. Moreover, such exchanges help create a community of scholarship around teaching that is based on a concrete, discipline-specific context.

**Decision making**

Accomplishments in teaching are becoming a more important factor in administrative decisions such as tenure, promotion, reappointment, and merit increases. The teaching portfolio enables faculty and departments to insure that an instructor's work in teaching is judged using multiple forms of evaluation, seen by multiple eyes. This is important, since no one perspective can accurately represent faculty teaching. For instance, students can evaluate certain aspects of teaching that focus on classroom interactions, such as organization, rapport, and ability to stimulate discussion. On the other hand, faculty colleagues are in a position to judge items that are beyond the expertise of students, such as how up-to-date material is, how well a course is integrated into the curriculum, etc.

Self-evaluation and reflection are also important, especially for providing a context for understanding data about teaching effectiveness. The portfolio as a whole gives individual faculty a sense of control over the evaluation process. In addition, departments that encourage faculty to
submit portfolios will need to have discussions about what, if any, documents will be required and what will be left up to the individual faculty; how long the document can (or should) be; and how much reflection is required. Such discussions provide a useful venue for creating a shared sense of what constitutes good teaching in a department.

Graduate student portfolios

Graduate students who apply for faculty positions commonly use portfolios because many colleges and universities now require job applicants to provide some proof of teaching experience. Graduate students are turning to the portfolio as a way of organizing their work in this area. Currently, the requirements vary widely among schools. Some require just a list of courses taught or a reflective statement on teaching, and some ask for specific items (such as proposed syllabi for certain types of courses, student ratings, demonstrations of commitment to undergraduate research, etc.). The earlier in their teaching careers that graduate students begin to think about their portfolios, the more chance they will have to retrieve the documents they find most representative of their accomplishments. Aside from its value for the job market, the portfolio often represents the first time graduate students have had the opportunity to reflect on their teaching, which they often find both challenging and rewarding.

An Alternative to the Teaching Portfolio: Course Portfolios

A variation on the teaching portfolio is a course portfolio. As the name implies, these documents focus on a specific course, with a special emphasis on student learning. A course portfolio, therefore, is analogous to a scholarly project. It includes sections on goals (intended student learning outcomes), methods (teaching approaches used to achieve outcomes), and results (evidence of student learning) for a specific course.

Moreover, it is the relationship or congruence among these elements that makes for effectiveness. We expect a research project to shed light on the questions and issues that shape it; we expect the methods used in carrying out the project to be congruent with the outcomes sought. And the same can be said of teaching.

By encompassing and connecting all three elements—planning, implementation, and results—the course portfolio has the distinctive advantage of representing the intellectual integrity of teaching. (Cerbin, 1993, p. 51)

Course portfolios offer advantages for the person developing them as well as for the curriculum. For the faculty member developing the portfolio, the advantages are similar to those of assembling a teaching portfolio (e.g., self-reflection and a chance to compare intentions with outcomes), but with more in-depth insight into the impact of teaching on students. For departments, course portfolios can provide continuity and reveal gaps in the curriculum. For example, a course portfolio becomes a record of the purpose and results of a course that can be passed on to the next person in charge of that course or to the faculty member who teaches the next course in a sequence. By examining a set of course portfolios, a curriculum committee can gain an overview of what students are learning and what is missing, which could help with the process of curriculum revision.

How are Portfolios Evaluated?

Just as there is no one model for a teaching portfolio, there is no one method for evaluation. Again, this is a strength of the portfolio, since it means that individual units will need to develop criteria for evaluation and make them relevant to faculty in that unit. The process of deciding on criteria can also help to clarify what faculty in that unit value with respect to teaching. For one example of an evaluation scheme, see Appendix B.

As units develop criteria for evaluating portfolios, they should first consider the ways they plan to use the portfolio. Will portfolios be limited to faculty being considered for tenure or promotion or for instructors nominated for teaching awards, or will all faculty prepare a course portfolio in preparation for a department-wide curriculum review? These purposes differ and so should the requirements for the portfolios involved.

Once the purpose is clear, faculty will probably want to create guidelines for assembling portfolios. While it is important to maintain the flexibility of the portfolio, it is also necessary to insure some degree of consistency in order to make evaluation fairer and more reliable.
Faculty might establish consensus on required items, such as a page limit for the overall size of the portfolio, the focus (a single course, an overview of teaching, or a combination), opportunities for reflection, or a template (so that faculty do not need to worry about format and can concentrate instead on the content). Ideally, such guidelines will be established with input from potential reviewers in the unit as well as those faculty who will be under review.

4. Fostering a culture of teaching

- Portfolios can provide a rich and contextualized source of evidence about teaching achievements that can be used for a variety of purposes, including evaluation, improvement, summary of faculty careers, and defining “good teaching” in a department.

How Can Faculty Get Started?

Faculty can begin at any time to collect materials for their portfolios. At first, this process might entail simply saving relevant materials related to teaching so that they are readily accessible for review. At some point the faculty member will need to sort through the materials and decide which ones best represent his or her teaching accomplishments. Often this process is enhanced when faculty collaborate with each other as they build their portfolios.

CRLT offers campus-wide workshops on teaching and course portfolios, and we can bring a customized workshop to departments. The focus of the workshop is to help faculty develop a clear idea of what a portfolio is and what items it might include and to give faculty an opportunity to begin a reflective statement on teaching. When workshops are conducted in a department, faculty can begin to answer the question, “What is good teaching in our department?” CRLT also provides one-on-one consultations for individual faculty who are working on their portfolios and for units as they develop a systematic approach to portfolios.

Advantages of Portfolios

In the AAHE monograph The Teaching Portfolio: Capturing the Scholarship of Teaching, the authors describe four main benefits of the teaching portfolio (Edgerton, Hutchings, & Quinlan, 1991, pp. 4-6). Course portfolios have similar attributes.

1. Capturing the complexity of teaching

- Portfolios contain evidence and reflection in the context of what is being taught to whom under what conditions.
- The portfolio can present a view of a teacher’s development over time.
- Entries in the portfolio can be annotated to explain their significance for the faculty member’s teaching.

2. Placing responsibility for evaluation in the hands of faculty

- Faculty are actively involved in presenting their own teaching accomplishments so that evaluation is not something done “to” them.
- Portfolios extend evaluation beyond student ratings and encourage peer review and collaboration.
- The need to evaluate portfolios can lead to discussions on standards for effective teaching.

3. Encouraging improvement and reflection

- Assembling a portfolio involves reflection.
- Because they involve reflection, portfolios allow faculty to compare their ideals with their actions, a first step in efforts to improve.
- A faculty member’s portfolio reveals both products (evidence) and processes (reflection) of teaching to colleagues who read it.

References


Possible items for inclusion

Faculty members should recognize which of the items which might be included in a teaching dossier would most effectively give a favorable impression of teaching competence and which might better be used for self-evaluation and improvement. The dossier should be compiled to make the best possible case for teaching effectiveness.

THE PRODUCTS OF GOOD TEACHING
1. Students’ scores on teacher-made or standardized tests, possibly before and after a course has been taken as evidence of learning.
2. Student laboratory workbooks and other kinds of workbooks or logs.
3. Student essays, creative work, and project or field-work reports.
4. Publications by students on course-related work.
5. A record of students who select and succeed in advanced courses of study in the field.
6. A record of students who elect another course with the same professor.
7. Evidence of effective supervision of Honors, Master’s or Ph.D. theses.
8. Setting up or running a successful internship program.
10. Documentary evidence of help given by the professor to students in securing employment.
11. Evidence of help given to colleagues on teaching improvement.

MATERIAL FROM ONESELF
Descriptive material on current and recent teaching responsibilities and practices.
12. List of course titles and numbers, unit values or credits, enrollments with brief elaboration.
13. List of course materials prepared for students.
14. Information on professor’s availability to students.
15. Report on identification of student difficulties and encouragement of student participation in courses or programs.
16. Description of how films, computers or other nonprint materials were used in teaching.
17. Steps taken to emphasize the interrelatedness and relevance of different kinds of learning.

Description of steps taken to evaluate and improve one’s teaching.
18. Maintaining a record of the changes resulting from self-evaluation.
19. Reading journals on improving teaching and attempting to implement acquired ideas.
20. Reviewing new teaching materials for possible application.
21. Exchanging course materials with a colleague from another institution.
22. Conducting research on one’s own teaching or course.
23. Becoming involved in an association or society concerned with the improvement of teaching and learning.
25. Using general support services such as the Education Resources Information Centre (ERIC) in improving one’s teaching.
26. Participating in seminars, workshops and professional meetings intended to improve teaching.
27. Participating in course or curriculum development.
28. Pursuing a line of research that contributes directly to teaching.
29. Preparing a textbook or other instructional materials.
30. Editing or contributing to a professional journal on teaching one’s subject.

INFORMATION FROM OTHERS
Students:
31. Student course and teaching evaluation data which suggest improvements or produce an overall rating of effectiveness or satisfaction.
32. Written comments from a student committee to evaluate courses and provide feedback.
33. Unstructured (and possibly unsolicited) written evaluations by students, including written comments on exams and letters received after a course has been completed.
34. Documented reports of satisfaction with out-of-class contacts.
35. Interview data collected from students after completion of a course.
36. Honors received from students, such as being elected “teacher of the year”.

Colleagues:
37. Statements from colleagues who have observed teaching either as members of a teaching team or as independent observers of a particular course, or who teach other sections of the same course.
38. Written comments from those who teach courses for which a particular course is a prerequisite.
39. Evaluation of contributions to course development and improvement.
40. Statements from colleagues from other institutions on such matters as how well students have been prepared for graduate studies.
41. Honors or recognition such as a distinguished teacher award or election to a committee on teaching.
42. Requests for advice or acknowledgement of advice received by a committee on teaching or similar body.

Other sources:
43. Statements about teaching achievements from administrators at one’s own institution or from other institutions.
44. Alumni ratings or other graduate feedback.
45. Comments from parents of students.
46. Reports from employers of students (e.g., in a work-study or “cooperative” program).
47. Invitations to teach for outside agencies.
48. Invitations to contribute to the teaching literature.
49. Other kinds of invitations based on one’s reputation as a teacher (for example, a media interview on a successful teaching innovation).

<table>
<thead>
<tr>
<th>QUESTION</th>
<th>DOSSIER MATERIALS</th>
<th>SUGGESTED FOCUS IN EXAMINING DOSSIER MATERIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What is the quality of materials used in teaching?</td>
<td>Course outline</td>
<td>Are these materials current?</td>
</tr>
<tr>
<td></td>
<td>Syllabus</td>
<td>Do they represent the best work in the field?</td>
</tr>
<tr>
<td></td>
<td>Reading list</td>
<td>Are they adequate and appropriate to course goals?</td>
</tr>
<tr>
<td></td>
<td>Text used</td>
<td>Do they represent superficial or thorough coverage of course content?</td>
</tr>
<tr>
<td></td>
<td>Study guide</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Description of non-print materials</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hand-outs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Problem sets</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assignments</td>
<td></td>
</tr>
<tr>
<td>Peer Reviewer's Rating:</td>
<td>Low ________ ____ ____ ____ ____ ____ ____ ____ ____ __________________________</td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. What kind of intellectual tasks were set by the teacher for the</td>
<td>Copies of graded examinations</td>
<td>What was the level of intellectual performance achieved by the students?</td>
</tr>
<tr>
<td>students (or did the teacher succeed in getting students to set for</td>
<td>Examples of graded research papers</td>
<td>What kind of work was given an A?  a B?  a C?</td>
</tr>
<tr>
<td>themselves). And how did the students perform?</td>
<td>Examples of teacher's feedback to students on written work</td>
<td>Did the students learn what the department curriculum expected for this course?</td>
</tr>
<tr>
<td></td>
<td>Grade distribution</td>
<td>How adequately do the tests or assignments represent the kinds of student performance specified in the course</td>
</tr>
<tr>
<td></td>
<td>Descriptions of student performances, e.g., class presentation, etc.</td>
<td>objectives?</td>
</tr>
<tr>
<td></td>
<td>Examples of completed assignments</td>
<td></td>
</tr>
<tr>
<td>Peer Reviewer's Rating:</td>
<td>Low ________ ____ ____ ____ ____ ____ ____ ____ ____ __________________________</td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. How knowledgeable is this faculty member in subjects taught?</td>
<td>Evidence in teaching materials</td>
<td>Has the instructor kept in thoughtful contact with developments in his or her field?</td>
</tr>
<tr>
<td></td>
<td>Record of attendance at regional or national meetings</td>
<td>Is there evidence of acquaintance with the ideas and findings of other scholars?</td>
</tr>
<tr>
<td></td>
<td>Record of colloquia or lectures given</td>
<td>(This question addresses the scholarship necessary to good teaching. It is not concerned with scholarly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>research publication.)</td>
</tr>
<tr>
<td>Peer Reviewer's Rating:</td>
<td>Low ________ ____ ____ ____ ____ ____ ____ ____ ____ __________________________</td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Has this faculty member assumed responsibilities related to the</td>
<td>Record of service on department curriculum committee, honors program, advising</td>
<td>Has he or she become a departmental or college citizen in regard to teaching responsibilities?</td>
</tr>
<tr>
<td>department's or University's teaching mission?</td>
<td>board of teaching support service, special committees (e.g., to examine grading</td>
<td>Does this faculty member recognize problems that hinder good teaching and does he or she take a responsible part</td>
</tr>
<tr>
<td></td>
<td>policies, admission standards, etc.)</td>
<td>in trying to solve them?</td>
</tr>
<tr>
<td></td>
<td>Description of activities in supervising graduate students learning to teach.</td>
<td>Is the involvement of the faculty member appropriate to his or her academic level? (e.g., assistant professors</td>
</tr>
<tr>
<td></td>
<td>Evidence of design of new courses.</td>
<td>may sometimes become over-involved to the detriment of their scholarly and teaching activities.)</td>
</tr>
<tr>
<td>Peer Reviewer's Rating:</td>
<td>Low ________ ____ ____ ____ ____ ____ ____ ____ ____ __________________________</td>
<td></td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. To what extent is this faculty member trying to achieve excellence</td>
<td>Factual statement of what activities the faculty member has engaged in to improve</td>
<td>Has he or she sought feedback about teaching quality, explored alternative teaching methods, made changes to</td>
</tr>
<tr>
<td>in teaching?</td>
<td>his or her teaching. Examples of questionnaires used for formative purposes.</td>
<td>increase student learning?</td>
</tr>
<tr>
<td></td>
<td>Examples of changes made on the basis of feedback.</td>
<td>Has he or she sought aid in trying new teaching ideas?</td>
</tr>
<tr>
<td>Peer Reviewer's Rating:</td>
<td>Low ________ ____ ____ ____ ____ ____ ____ ____ ____ __________________________</td>
<td>Has he or she developed special teaching materials or participated in cooperative efforts aimed at upgrading</td>
</tr>
<tr>
<td>Comments</td>
<td></td>
<td>teaching quality?</td>
</tr>
<tr>
<td>Reprinted by permission.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GF LAZOVIK 1979
UNIVERSITY OF PITTSBURGH

Peer Reviewer's Signature __________________________
Date __________________________

Appendix B

7
Appendix E: University of Michigan Lecturers' Employee Organization Contract

**Definition of Titles**

**Intermittent Lecturer** - instructional faculty who teach one course less than every semester.

**Adjunct Series** - instructional faculty who are employed elsewhere in the University at more than 50% capacity.

**Lecturer Series (I, II, II, IV)** - Lecturer I and II and Lecturer II and IV are parallel tracks. I and II are primarily teaching assignments; III and IV involve other duties ("additional ongoing administrative or service duties within the academic unit, and/or requires a range of instructional expertise").

**Performance Review**

**Lecturer I**: interim review no later than the end of the fifth semester; major review after eight consecutive F/S semesters.

**Lecturer II**: first major review after eight consecutive semesters as a Lecturer I; if successful, instructor receives three-year appointment as L II. Second major review occurs in final year of new appointment; if successful, instructor receives an appointment for an additional five-seven years.

**Lecturer III**: interim review no later than the end of the fifth semester; if successful, promotion to L IV.

**Lecturer IV**: initial major review as L III results in three-year appointment as L IV; second major review in the final year of appointment; if successful, this results in a five-year appointment as L IV.

**Adjunct Series**: may undergo performance review upon written request; not mandatory.

**Intermittent Lecturer**: may undergo performance review upon written request; not mandatory.

(Effective March 21, 2013, through April 20, 2018)
Chapter 4: The role of student surveys in the evaluation of teaching

A true story\(^1\): Once upon a time, a dean asked the manager of a university's student instructional survey to find the highest-rated teacher at the university. The manager patiently explained that it would make no sense to use student instructional surveys in this way – that student ratings are affected by so many factors that it makes no sense to treat them as pure measures of instructor quality. The dean insisted. The result: the top-rated instructor in the entire university turned out to be a faculty member who taught an upper-level humanities course with an enrollment of five students.

It is not hard to see why this isn't necessarily an appropriate way to identify outstanding teachers – and why the evaluation of teaching requires far more than a single number. However, as we note below, student instructional rating surveys have the potential to provide useful information, provided they are properly designed, provided the results are carefully interpreted, and provided the results are considered along with other assessments of teaching (including teaching portfolios, peer review of curricular materials and the like).

At the outset, we note that student surveys can potentially be (and often are) used in at least two ways: first, in personnel decisions affecting instructors (e.g., concerning promotion and tenure); and, second, as a means of identifying aspects of an instructor's teaching that either deserve commendation or need improvement. The first of these requires a certain amount of uniformity across disciplines and departments; the second requires a certain amount of flexibility across disciplines and departments. In each case, however, reliance on student surveys to the near-total exclusion of virtually every other method of assessing teaching is seriously misguided and inappropriate, and will not be adequate for a robust evaluation of the quality of instruction or student learning. The true story noted in our opening paragraph is a very serious cautionary tale.

A. Limitations of student ratings surveys

The limitations of student ratings surveys are well-known and much-discussed. The response rate to such surveys is usually 50 percent or less (sometimes much less).\(^2\) If (as is often the case) students are not required to respond to such surveys, then, of course, response is optional, and the respondents will therefore be self-selected.\(^3\)

A second problem is that, as some research\(^4\) has found, student ratings at a variety of universities tend to be related to discipline (teachers fare less well in the sciences than the

---

\(^1\) Some details have been changed to protect the innocent (and spare the guilty).
\(^2\) However, it should be noted that biased or misleading survey results arise not from a low response rate *per se* but, rather, from self-selection and other kinds of nonrandom response. On the other hand, when actual enrollment in a class is small, a low response rate will obviously entail too few actual responses to be meaningful.
\(^3\) Folk wisdom (not usually supported by actual evidence) holds that, when response is optional, the students who actually respond tend to be either highly favorable or highly unfavorable towards the instructor and the course, with much less representation of students with more moderate views. This does not necessarily mean that the average of all responses will reflect the sentiment of all students in the course, but there is no obvious basis for thinking that this will necessarily be the case. Still more problematic is that there is no obvious way to know how self-selection will tilt the average student rating relative to the sentiment of all students in the course.
\(^4\) See the Appendix to this chapter for a review of some of this research.
humanities), course level (enrollment tends to be self-selected, and more favorably disposed to the instructor, in upper-level relative to lower-level courses), class size (high-enrollment courses fare less well than low-enrollment courses), race, and gender. Thus, student ratings may reveal more about these other factors, and less about the quality of instruction per se, than one might imagine. Considered as the major or sole means of identifying outstanding teachers, or of measuring the quality of an instructor's teaching, or of measuring what students have actually learned, student ratings may be highly unsatisfactory.

Leaving aside problems about response rate, self-selection and the influence of extraneous factors, the information in student surveys may be less helpful than it could be. As a case in point, in Appendix A, we reproduce the University's standard instructional ratings survey form. Responses to some of the questions on this form tend to be highly correlated with other responses (e.g., quality of course vs. quality of instruction (questions 9 and 10 on the form)), so that a large number of questions may not generate much additional information beyond what would be obtained from a survey with fewer, better-focused questions. On the other hand, student surveys such as this one may not provide information on other topics that may be of considerable importance, e.g., the extent of cheating or the expected course grade.

Finally, (and, again, leaving aside all of the limitations noted above), it is not at all obvious what student ratings actually measure. Are students actually well-equipped to judge the quality of instruction, and of their instructors? Does a low rating for a course suggest that it has not been well-designed or that the readings and other course materials are inadequate? Does a low rating for an instructor mean that the instructor is incompetent? Or, alternatively, does a low rating merely mean that the course material is complex and requires extra effort, or that grading is rigorous? Does a low rating for a course and/or its instructor mean that students learned little? Does a high rating mean that students learned a great deal? There are no clear answers to these questions – particularly when student ratings for a course and/or instructor are considered in isolation from other important information about the course and/or the instructor.

These concerns about student ratings are serious. However, properly considered, student ratings do have the potential to provide some useful information (even if the extent to which this is true can easily be overstated). For example, as Stark and Freishtat (2014) note, even though “students cannot rate [teaching] effectiveness,” it is nevertheless true that “students are in a good position to observe some aspects of teaching, such as clarity, pace, legibility, audibility, and their own excitement (or boredom).” Thus, as we now explain, we believe that concerns about student ratings can be adequately addressed, and that such ratings can provide useful information, provided the underlying ratings surveys are properly designed, and provided the ratings are carefully interpreted in conjunction with other information about the courses and the instructors. As Stark and Freishtat (2014) comment, “Student ratings of teaching are valuable when they ask the right questions, report response rates and score distributions, and are balanced by a variety of other sources and methods to evaluate teaching.”

---

5 A professor relates that, during a question-and-answer session after another professor's lecture on international-security affairs, an undergraduate "waved away the professor's views, saying, 'Well, your guess is as good as mine.' 'No, no, no,' the professor said emphatically. 'My guesses are much, much better than yours."" (See Tom Nichols, "Our Graduates Are Rubes," The Chronicle Review, Jan. 20, 2017, p. B3.)
2. **Better design and interpretation of student ratings surveys**

Experience at various universities suggests that, except in unusual circumstances, it is difficult to achieve a response rate to online student ratings surveys that is substantially in excess of 50 percent. These surveys are given towards the end of the semester, when students are particularly busy with coursework; and students often believe (not without reason!) that they won't benefit much if at all if they respond to such surveys. It should be kept in mind that a student taking four or five classes will receive four or five different surveys. It would not be at all surprising if “survey fatigue” helped to reduce the rate of response to some or even all of the surveys.

Making it mandatory for students to respond – making such response a course requirement, on a par with taking exams and completing other course requirements – will certainly increase the response rate, and may also reduce the extent to which respondents are self-selected. This will require both (a) mechanisms for identifying students who fail to respond that preserve the anonymity of all student responses and (b) appropriate sanctions for students who fail to respond. We recommend that the University develop a mechanism that can identify (and prod) students who fail to respond to online requests for student instructional ratings. Sanctions for non-response, which might best be developed on a departmental basis, might include withholding grade information or transcripts.

However, relying on compulsion as a means of increasing survey response rates may well backfire: students who are annoyed by repeated reminders, and by the knowledge that they have to devote time to responding to multiple surveys in order to get their transcripts, may simply enter responses randomly (or enter only the most negative or positive response). In this way, compulsion may result in a higher response rate only by reducing the quality of the information obtained.

Rather than rely on compulsion, it may be preferable to encourage response by developing carrots (and incentives) as well as sticks – and by emphasizing to students that student ratings surveys do give them a voice and an organized way to participate in the assessment of teaching. Some units – e.g., the Byrne Seminars – conduct their surveys the old-fashioned way, i.e., by using paper forms. A survey of this kind can be administered during the final class of the semester or at the conclusion of the final exam for the class. In such circumstances, virtually all of the students in the class will respond.

We previously noted that, according to some research studies of other universities, student ratings of courses tend to be related to academic discipline, course level, class size, and instructor characteristics such as race and gender. Contrary to popular opinion, the research literature is by no means unanimous in supporting these conclusions, particularly with respect to race and gender. We strongly recommend that the University undertake a systematic study of how, at Rutgers, these and other factors are related to student ratings.

We would also note that the means for obtaining more, and different, information from student surveys already exists, and can be used to a much greater extent than is the case at present. In particular, individual departments that wish to design their own student instructional
ratings surveys already have the means to do so (although, in our experience, these means are sadly underused). Such surveys could (e.g.) ask students what grades they expect to receive, about the extent of cheating, etc. Individual departments can easily (re)design student surveys in response to their specific needs.

Indeed, some individual departments have already done so. Departments have always been able to add additional questions of their own design on the paper survey forms and many did so, e.g., the departments of Physics, English, and Music in New Brunswick and the Department of Economics in Newark. Departments can still easily add additional questions to the online SIRS if they inform CTAAR about the additional questions before the survey period begins. Faculty can now add their own questions to the online SIRS in the Sakai survey system. Newer systems such as EvaluationKit give departments and individual faculty even more flexibility to add their own questions.

We also believe that departments can, and should, pay more attention to the details of the results of student surveys. Careful design – and analysis – of ratings surveys has the potential to provide much useful information. In particular, individual instructors and departments can obtain the (anonymous) unit record data on students’ responses and analyze them. For example, one of our own task force members designed his own anonymous in-class survey independent of the SIRS. This survey included questions that do not appear on the standard student instructional ratings survey form, e.g., about the expected course grade and whether the respondent had also taken the online survey. Over 95% of students in his classes responded (the survey was administered immediately at the end of the final exam). This instructor found (e.g.) that students who said they had a “strong prior interest” in their course rated their instructor much more favorably than students who did not have a strong prior interest; and that there was little difference in the ratings of students who said they did or did not complete the regular online instructional rating survey.

Perhaps the most important point about student ratings surveys is that the results of these surveys will be most informative when – but only when – they are considered in conjunction with substantial additional information about the course and instructor to which they refer. As we noted in previous chapters, this additional information can, and should, be derived in various ways, e.g., from each instructor's teaching portfolio, classroom visits, and the like. This additional material will allow the results to be placed in context, and to permit a much better-informed assessment of teaching.

Unfortunately, however, in our experience, all too often the responses that get most if not all of the attention are the responses to Questions 9 and 10 on the “standard form” in Appendix B, on the quality of the course and the quality of the instructor. This is particularly the case because, in general, departments are required to provide student responses to these two questions on the University's Form 1A, which is part of the documentation considered in all decisions on promotion. (See Appendix C, which reproduces the guidelines concerning student course ratings

---

6 It should be noted that departments are able to request additions to and modified versions of the standard questions in SIRS.
7 Perhaps not surprisingly, the instructor also found that students who said they expected a higher grade rated their instructor more highly than did students who said they expected a lower grade.
on Form 1A.) On the other hand, and despite all the above-noted limitations and concerns about student ratings, the responses to these two questions are often the only information about teaching that departments are required to provide in connection with promotion decisions. Strictly speaking, this is a defect not in the student ratings surveys but, rather, in the way in which they are used; we say more about this in the following chapter.
Appendix A: A brief review of the research literature on student evaluations of teaching

The research literature on student evaluations of teaching is enormous and growing.\(^8\) We limit ourselves here to a very brief discussion of two kinds of studies: studies that consider whether student evaluations of teaching provide useful information about teaching effectiveness; and studies that consider whether student evaluations of teaching are affected by gender bias.

**Student evaluations of teaching and teaching effectiveness**

Do student instructional ratings provide useful information about teaching effectiveness? One of the most sophisticated studies of this question is by Carrell and West (2010). They found that in introductory science courses, students of junior professors rated them more highly and got higher grades than did students of senior professors – but students who had senior professors for their introductory courses actually did better in follow-on courses than the students who had junior professors for their introductory instructors. Carrell and West conclude that the junior professors adopted a cookie-cutter approach that allowed students to do well during the introductory course, but did not equip them to understand the complexity of the subject matter. In contrast, the senior professors adopted a more rigorous approach; students didn't like it, but they were better prepared for the follow-on course.

Unfortunately, however, student evaluations of teaching may also be influenced by factors other than teaching effectiveness – so that such evaluations may not be a reliable or unambiguous measure of teaching effectiveness per se. For example, Hamermesh and Parker (2005) found that, for both male and female instructors, the instructor's physical appearance was statistically significantly related to evaluations given to the instructor by his or her students. Similarly, Ambady and Rosenthal (1993) found that student evaluations of teaching could be predicted using students’ reaction to a 30 second silent video of the instructor.

However, the case against student evaluations as measures of teaching effectiveness is by no means decisive. Beleche et al. (2012) find that, after controlling for a substantial number of covariates, students’ course evaluation ratings are positively and statistically significantly related to an objectively-graded post-test measure of student learning. Based on a meta-analysis of research studies, Clayson (2009) concludes that there appears to be “a small average positive association” between student learning and student evaluations of teaching, but that the relationship is not universal.

**Student evaluations of teaching and gender bias**

Are student evaluations of teaching affected by gender bias? Here, the research literature offers evidence on every side of the question.

A number of analyses have found that female instructors receive lower student evaluation ratings than do male instructors, even when other factors are taken into account. For example, Hamermesh and Parker (2005) found that students’ ratings of female faculty were lower than

---

\(^8\) For reviews of some of the older research literature on this subject, see Marsh (1984) and Marsh and Bailey (1993). For a review covering more recent research, see Stark and Freishtat (2014).
ratings of male faculty who were the same in terms of race, tenure track status, physical attractiveness, and other factors.

In contrast, Basow (1995, p. 656) found that “[o]verall, the ratings of male professors appeared to be unaffected by student gender. In contrast, female professors tended to receive their highest ratings from female students and their lowest ratings from male students.” On the other hand, Cramer and Alexitch (2000, p. 157) found that “students rated female professors more positively than male professors”; whereas Young et al. (2013, p. 290) found that “all students evaluated female STEM professors as more favorable role models than male professors.” (For a similar finding, see Centra and Gaubatz, 2000.) Relying on a meta-analysis of 39 studies of student teacher ratings, Feldman (1993, p. 153) concluded that “female teachers received a slightly higher overall evaluation than did the male teachers.” He added, however, that although this correlation was statistically significant, it was “so small as to have virtually no practical significance.”

Studying sex differences in instructor ratings is, inevitably, a difficult enterprise: Since student ratings vary with a variety of factors (e.g., the size of the class, the academic discipline of the course, the sex composition of the students in the course, and even the time of day and day(s) of the week when the course meets), it is hard to isolate sex differences in students’ ratings from other, confounding factors. This may help explain why results concerning sex differences in ratings of instructors vary from study to study, and why it is hazardous to attempt to generalize from a single study, or even a number of studies.

**References for Appendix A**


Appendix B: Standard online student ratings survey questions

1. The instructor was prepared for class and presented the material in an organized manner.

2. The instructor responded effectively to student comments and questions.

3. The instructor generated interest in the course material.

4. The instructor had a positive attitude toward assisting all students in understanding course material.

5. The instructor assigned grades fairly.


7. I learned a great deal in this course.

8. I had a strong prior interest in the subject matter and wanted to take this course.

9. I rate the teaching effectiveness of the instructor as:

10. I rate the overall quality of the course as:

11. What do you like best about this course?

12. If you were teaching this course, what would you do differently?

13. In what ways, if any, has this course or the instructor encouraged your intellectual growth and progress?

14. Other comments or suggestions.
Appendix C: Instructions on use of student ratings survey data in Form 1A

For each course for which summary student evaluation data are available, include the number of student evaluation responses received, and the instructor and departmental mean values for questions 9 and 10 on the University's Student Instructional Rating Form. If units use a different rating form, please indicate maximum rating value. If evaluations are not included for a specific course, please account for missing evaluations.

Example:

<table>
<thead>
<tr>
<th>- COURSE INFORMATION -</th>
<th>- COURSE EVALUATION -</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evaluating Responses</td>
</tr>
<tr>
<td>S/ Yr</td>
<td>Number</td>
</tr>
<tr>
<td></td>
<td>Course Title</td>
</tr>
<tr>
<td>E. F1</td>
<td>161:111:11</td>
</tr>
<tr>
<td>Women’s Studies</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 5: What’s next?

Improvement in the methods used in evaluation of teaching at Rutgers—New Brunswick (RU-NB) has been discussed for many years. Unfortunately, very little about the evaluation of teaching at RU-NB has actually changed. In what follows, we discuss our recommendations concerning the organizational and administrative steps that need to be taken in order to make better evaluation of teaching a reality. The crucial requirement for the kind of change and improvement that we hope for is leadership – and strong emphasis on getting the job done – on the part of the administration.

Based on the discussion in the foregoing chapters, we first offer several recommendations for improving the evaluation of teaching at RU-NB. We then discuss how these recommendations might actually be implemented and who might be made responsible for their implementation.

A. **Recommendations for better evaluation of teaching**

As noted in Chapter 4, research on student course rating surveys offers few unambiguous conclusions. **We therefore recommend that the RU-NB Chancellor should commission studies based on data derived from the Student Instructional Ratings System (SIRS).** Such studies might consider, for example, whether (other things being equal) student responses about “course quality” vary by characteristics of the course or of the instructor (e.g., academic discipline; course enrollment; instructor gender, rank and faculty status). Further studies might examine the impact of changes in the evaluation system.

**We recommend that the Chancellor should establish a Teaching Evaluation Council (TEC) for RU-NB,** chaired by the Vice Chancellor for Undergraduate Academic Affairs. The Council can be a subset of, or be subsumed as, the Vice Chancellor’s current Council of Associate Deans. The TEC would serve as a clearinghouse for ideas and suggestions on the improvement of the evaluation of teaching; and would act as an advocate, through the Vice Chancellor, for resources for improving teaching, its assessment, and the evaluation of teaching in the promotion process for the instructors, full time and part time, tenure-track and non-tenure track, throughout RU-NB. In addition, it would assist departments as they prepare or revise their individual plans and procedures for evaluating teaching.

**We recommend that each department at RU-NB should be required to prepare a systematic plan for teaching evaluation,** drawing on (e.g.) the practices and techniques summarized in Chapters 2 and 3 and on the assistance of the TEC. **We further recommend that each department submit its teaching evaluation plan to the dean of its school or college for review and, if necessary, revision.**

Given the very diverse nature of the academic departments at RU-NB, one size – and one plan – will certainly not fit all departments, and it would make no sense to require that every department must conduct its evaluation of teaching in the same way. It is, however, essential that each department actually have a thorough and comprehensive plan, covering each of the various categories of faculty (TT, NTT, PTL, etc.) who are part of its instructional effort. For
example, a plan for evaluating an individual faculty member’s teaching might consist of (a) a teaching portfolio and teaching statement prepared by the faculty member; (b) one or more visits to the faculty member’s classes by one or more senior colleagues; (c) a report by these senior colleagues on the faculty member’s teaching; and (d) one or more tabulations derived from one or more student instructional ratings, accompanied by a written commentary prepared by senior colleagues. The plan would specify how frequently each aspect of the evaluation would take place for each kind of faculty member; for example, for TT faculty, classroom visits might take place prior to consideration for (a) reappointment, (b) promotion with tenure, (c) subsequent promotion and (d) periodic post-tenure review and/or rank review.

Although individual departments should be able to devise the particular combination of teaching evaluation procedures and techniques that will best meet their needs, they should all be required to produce an overall summary assessment of each faculty member’s teaching, based on the entire set of procedures and techniques used in its evaluation (e.g., classroom visits, evaluation of the teaching portfolio, etc.). In particular, we recommend that the evaluation of each faculty member’s teaching should use a categorical rubric and an indication of where the faculty member fits in that rubric: e.g., “significantly exceeds expectations,” “meets expectations,” “falls short of expectations,” and “significantly falls short of expectations.” It should be noted that, when it develops its teaching evaluation plan, each department should clearly describe what its expectations for teaching actually are, and what would constitute falling short or exceeding those expectations.

B. Implementing procedures for teaching evaluation

As noted in Chapter 1 and at the beginning of the present chapter, teaching evaluation is a bit like the weather: everyone talks about it, but in all too many cases, disappointingly little is done about it. In what follows, we outline the steps that must be taken if better evaluation of teaching is to become a reality.

First, as noted in the previous section, the dean of each RU-NB school or college (e.g., School of Arts & Sciences, School of Business, etc.) will request each of the unit’s departments to prepare a written policy statement describing, in detail, their policies and procedures for teaching evaluation. Each dean will then review the individual departmental statements, requesting revisions as needed, and seeking input as appropriate from the Teaching Evaluation Council. Finally, as noted below, each dean will report to the Vice Chancellor for Undergraduate Academic Affairs on teaching-evaluation plans for each department.

As determined by the Chancellor and Vice-Chancellor, the TEC will issue a request annually or biannually to each school for a report on the methods and programs used to evaluate teaching within the departments within each of its departments. The reports will be expected to respond to four basic questions about how teaching is evaluated in each department within each school. The four basic areas are:

\footnote{These evaluations would not necessarily rely solely on the existing set of questions in SIRS. As noted in Chapter 4, departments have considerable leeway in designing their own surveys and/or adding questions to the SIRS survey, and will certainly be able to do so in order to meet their needs.}
Use of the student course ratings surveys: How are the data drawn from SIRS and/or other surveys of students used in the evaluation of teaching in each department and program? What changes would each department and program desire to make SIRS and/or these other surveys more useful?

Each instructor is expected to create a teaching portfolio. In the reports sent to the TEC by each school, an explanation must be given as to how teaching portfolios are used to evaluate teaching and in the promotion process.

Each department and program is expected to create a process of peer review, including a departmental review of teaching materials in each course and for each instructor, and an in class visit by a team of peers at least once every three years. In the reports prepared by the schools for the TEC, each school will describe how peer review is conducted, and how its results are used to evaluate teaching.

Service courses and cross-disciplinary instruction: If a school has programs that cross-list courses, or provides service courses for other schools, then the school should report on the mechanisms to evaluate teaching in these circumstances through the cooperation of the different programs and schools involved.

The TEC will write a brief narrative response to each school, sharing good ideas for the evaluation of teaching across the schools, and making suggestions for improvement. The Vice Chancellor will coordinate the writing of these narratives and will return them to each respective school within three months of receiving the annual reports.

Upon completion of the narratives, the Vice Chancellor will meet with the Chancellor to present the narratives and a brief synopsis of the state of the evaluation of teaching in New Brunswick.

Using the TEC Narrative Responses as a foundation, the Chancellor will consider how teaching evaluation is conducted in each school, and how it is progressing over time. A school’s progress in implementing its new procedures for teaching evaluation should be an important element of the Chancellor’s review, for merit pay purposes, of the performance of the school’s dean.

C. Concluding comments

As noted above, the past twenty years or so have seen numerous recommendations from numerous groups concerning improvement of the University’s evaluation of teaching. Unfortunately, although many of these recommendations have been discussed, not many have actually been implemented.2

2 The 2002 Senate Resolution on Charge S-0109 Best Practice in Assessment of Teaching is one case in point. A more recent example is the brief but cogent New Brunswick Faculty Council Resolution on Evaluation of Teaching of 2007. Ideas are presented and recommendations made, usually the same or similar ones again and again, and yet there is no follow-through.
The importance of establishing a structure of accountability for implementing and maintaining an enhanced system of evaluating teaching is self-evident. Without one, nothing will change, as we have seen in the past. This structure must be sustainable. It must be collaborative and collegial, with information and ideas moving from the departments and deans to the Chancellor and back again in a coherent and transparent manner. Most of all, supporting it should be a top priority for University administrators in New Brunswick at every level, from the Chancellor and Provost to the deans and department chairs.

It may be objected that implementing the recommendations we have made will require considerable time, effort, and resources. That is indeed the case. At issue is whether the University collectively believes that good teaching is important and is a high priority. If so, then, however challenging it may be, evaluating good teaching is eminently worth significant investments of time, effort, and resources. We believe that the University should no longer postpone making these investments. We believe that the time to act has now come.