
Report of the Task Force on Evaluating Teaching

October 2018



2017-2018 Task Force on Evaluating Teaching

Chair:

Gillian Pierce, Assistant Provost for Academic Assessment, Office of the Provost

Members:

Suzanne Chapin, Professor of Mathematics Education, School of Education

Terry Everson, Associate Professor of Music, College of Fine Arts

Ray Fisman, Slater Family Professor in Behavioral Economics, College of Arts & Sciences

Christopher Gill, Associate Professor of Global Health, School of Public Health

Daryl Healea, Assistant Dean for Curriculum & Enrollment Services, College of Arts & Sciences

Roland Jaeckel, Director, Educational Technology, Office of Digital Learning & Innovation

Deborah Jaramillo, Associate Professor of Film & Television Studies, College of Communication

Rebecca Kinrade, Master Lecturer, Writing Program, College of Arts & Sciences

Chris Paal, Assistant Vice President and University Registrar

Michael Smith, Associate Professor of Biomedical Engineering, College of Engineering

Matthew Trevett-Smith, Director, Center for Teaching & Learning

Staff:

Zachary Crawford, Administrative Coordinator, Office of the Provost

Executive Summary

In May 2017, the Task Force on Evaluating Teaching was charged with gathering information about the language and content of student course evaluations at Boston University and considering how this content aligns with current research and national conversations about how best to evaluate university teaching. The Task Force was further charged with drafting a University-wide course feedback form with both common content and questions and flexible options for schools, colleges, programs and departments to tailor the surveys to their specific needs. Additionally, the Task Force was asked to make recommendations about a move to collect student course feedback online. Finally, the Task Force was invited to comment on the place of student course feedback in the overall process of evaluating teaching on campus, considering the role of other measures, such as peer evaluation of teaching, self-reflection, and review of portfolios of teaching materials, all with the goal of creating a set of standard guidelines and practices for evaluating teaching on the Charles River and Medical Campuses.

After extensive research in the literature on evaluating teaching, study of practices at peer institutions, and broad consultation with students, faculty, and other stakeholders at Boston University, the Task Force makes the following key recommendations:

1. Boston University should adopt a university-wide online student feedback survey that is brief and includes both quantitative and qualitative elements. In addition to the core questions given to every School and College, additional, customizable questions can be included at the discretion of individual Schools, Departments, and faculty members.
2. Boston University should adopt online course feedback forms following set guidelines for administering the evaluations to ensure maximum response rates (generally, they should be administered in class during a predetermined two-week window to maintain response rates comparable to paper evaluations). A move to provide course feedback online ensures that all stakeholders receive results quickly and efficiently.
3. The University should require multiple measures in order to responsibly evaluate teaching, including peer evaluation and teaching portfolios, for all personnel decisions, including contract renewal.

4. Midterm course feedback surveys should be made available to all faculty in all courses, ideally through the same online mechanism as end-of-semester surveys. The emphasis of midterm course evaluation is entirely formative, so questions should be customizable, and results would be viewable only by the faculty member teaching the course. (See Appendix D for a sample survey.)
5. Results of end-of-semester feedback surveys should be disclosed to the student body. We recommend that comments in response to a single, open-ended question on the standard course feedback form be shared with students along with a summary of quantitative results. Further, there should be a process by which faculty or staff can redact student comments to remove inappropriate language or content. Students should also have the opportunity to share course information through a student-created web page and have access to published course syllabi to use in considering course selection.
6. Boston University should continue to keep pace with initiatives at AAU peer institutions to better evaluate teaching and emphasize its value in tenure and promotion processes and other faculty reward systems.

The Task Force believes that these steps are necessary to promote excellent teaching at Boston University and to remain aligned with our local and AAU peers.

Introduction

The Boston University Task Force on Evaluating Teaching began its work to draft a new, university-wide instrument to collect student course feedback amid mounting concerns about what standard student feedback surveys do and do not measure. Frequent articles and opinion pieces about bias in course evaluations and the need for better measures surfaced weekly as the Task Force met. At the same time, a newly-released Wall Street Journal/NBC news poll revealed increasing public skepticism about the value of a college degree, with the majority of young adults responding that college is not worth the cost.¹ The current national conversation and concerns about the quality and value of undergraduate education—and the loss of trust in higher education expressed by many, from across the political spectrum—relates directly to the quality of teaching on university campuses and impels us to search for the best possible

¹ Wall Street Journal/NBC news Sept. 7, 2017.

measures for evaluating and improving it. As a tuition-driven institution, Boston University must demonstrate the value of a BU degree to its students, who are seeking an intellectually rich and stimulating environment and challenge in the classroom from faculty who are dedicated to teaching. One essential way for universities to reclaim the narrative about the value of the degrees they grant is to recognize the importance of teaching in tangible ways, by seeking better evaluation tools to provide evidence of effectiveness. We can also strive for greater transparency about our criteria for excellent teaching through the development of common tools and resources and share results with students, who can use the information to make informed course selections.

Excellent teaching is at the heart of the educational experience Boston University offers as a residential research institution. As peer institutions and professional organizations such as the Association of American Universities (AAU) advance a national dialogue about ways to support and evaluate teaching and ensure that excellent teaching is appropriately recognized and rewarded on university campuses, it is critically important for Boston University to examine its approach to student course feedback and the place of student course feedback forms in the overall evaluation of both undergraduate and graduate teaching and, crucially, to develop multiple measures of teaching effectiveness to complement student feedback. In reviewing the literature and best practices of peer schools, it is clear that high quality teaching is best supported by nuanced, multi-factorial systems of assessing teaching of different types by feedback from a range of constituencies, including peers, students, faculty mentors, and departmental academic leadership.

As we consider how best to evaluate teaching, we must recognize the size of Boston University and the diversity of undergraduate and graduate teaching that takes place in its Schools and Colleges and campuses. Classes range from large lectures to small seminars and one-on-one instruction. Subject matter spans general education coursework, licensure, and professional preparation. Some have lab, project-based, or service-learning components, and some are team-taught. Courses are taught in face-to-face, online, and hybrid environments. Any approach to evaluating teaching must recognize the need to reflect all of the various kinds of teaching taking place on campus and the needs of individual schools and colleges.

A systematic approach to providing feedback about teaching is intended to support the needs of faculty members at various career stages. Graduate student teachers and early

career faculty would benefit from additional support, mentorship, and low-stakes opportunities to engage in substantive discussions about their teaching. Faculty with more experience teaching may benefit from feedback about specific teaching practices, teaching materials, or new approaches. Regular formative feedback through midterm course evaluations, peer observation, and peer review of teaching materials complements more high-stakes summative feedback typically required for promotion, tenure, merit review, and contract renewal, and generally contributes to improving quality of teaching within the research university by making conversations about teaching more frequent, more public, and more collaborative.

In response to its charge, the BU Task Force on Evaluating Teaching gathered information about the language and content of current BU student course evaluations. To begin, the Task Force conducted a twenty-five-question survey of Deans, Department Chairs, and Program Directors to better understand needs and current practices on the Charles River and Medical campuses. Predictably, each School and College has its own approach and a separate course evaluation form (or forms). Some Schools are using online course evaluations and have done so for many years; others use paper evaluations. We collected and reviewed all current student course evaluation forms on the Charles River and Medical Campuses to create a master inventory of current question types, and considered how our practices align with current research and national conversations about how best to evaluate teaching. Both the survey results and the inventory of course evaluation questions suggest that we currently lack a clear, shared definition of multiple measures of teaching effectiveness. Further, the quality of information about teaching from the instructor, students, and peers has been highly variable across the University, making it difficult to fairly and adequately assess teaching effectiveness and support improved teaching.

The Task Force was charged with drafting a University-wide student course feedback form to be used in both undergraduate and graduate courses, with a core set of common content and questions and a flexible section for Schools, Colleges, Programs, Departments, and individual faculty members to tailor the evaluations to their specific needs. We approached this task with some caution, as the research on the validity of student course evaluations is inconsistent. What is the correlation between strong teaching evaluations and student learning?² What do the evaluations really measure, and how susceptible are they to gender or

² Several large-sample research studies have in fact demonstrated a negative correlation between scores on student ratings of instruction and measurements of learning as demonstrated in subsequent courses (Weinberg, Hashimoto, and Fleisher, 2009, and Yunker and Yunker, 2003). The University of Oregon

racial bias? The 2017 AAU report, “Aligning Practice to Policies,” asserts that at research universities, “teaching effectiveness is overwhelmingly assessed using student evaluation surveys completed at the end of each course, despite evidence that these evaluations rarely measure teaching effectiveness [...], contain known biases [...], promote the status quo, and in some cases reward poor teaching,” and further that “the ease with which these surveys are applied has led to widespread misuse” (AAU 2017, 2), meaning that quantitative scores may be used indiscriminately to distinguish among teachers or influence hiring decisions, despite the lack of clear correlation between such scores and teaching effectiveness.

In the face of such concerns, the Task Force focused on drafting a set of questions that would provide meaningful feedback to faculty and emphasize those elements of a course that students are best qualified to describe and offer insight about, including their perceptions of their learning. To develop this instrument, we consulted numerous examples, particularly those from AAU institutions that have recently gone through a similar revision process³. Our own School of Public Health was in the middle of revising its student feedback survey when the Task Force began its work, and we were informed by their research and the instrument they developed.⁴

Since we were charged with making recommendations about a move to online course evaluations, the Task Force monitored the progress of a large-scale online course evaluation pilot in CAS/GRS, which further expanded to include COM, CFA’s SVA, MET, Wheelock, and LAW, and sought feedback from Schools and Colleges that have been using online course evaluations exclusively for some time (ENG, COM, Questrom, and the medical campus) to inform recommendations about a general move to online course evaluations at Boston University. Our research and review of literature showed that if relatively simple steps are taken, response rates need not differ between student surveys administered on paper and online course evaluations. Further, online course evaluations are cost effective and environmentally sustainable. Perhaps more importantly, they allow faculty and administrators to receive results quickly and easily, so that feedback can be better used to improve teaching.

Lastly, members of the Task Force consulted with graduate and undergraduate student groups and surveyed Department Chairs and Program Directors about if, when, and how results

recently conducted one such study of its own population, with similar findings that students who were taught by highly rated professors in prerequisite courses performed more poorly in follow-up courses.

³ The comprehensive Stanford University (2013), UC Berkeley (2009), Vanderbilt (2014), and Brown University (2018) reports were all instructive in this context.

⁴ The new SPH student feedback form is being piloted for validity in Fall 2018. The results of this pilot will help inform implementation of a new university-wide feedback form.

of student course feedback should be made available to students. We also looked at a range of practices in place at neighboring and peer institutions. While sharing results with students and other faculty members is standard in some Schools and Colleges, there is some resistance to the proposal in others. In light of this, and recognizing the significant pressure from students to provide additional reliable information about teaching in the courses they are considering, the Task Force grappled with two questions: which content from the proposed student survey would be most valuable to students in selecting courses, and how could we ensure a process of sharing those results that would meet with faculty approval? Making course syllabi publicly available would certainly also help students make informed decisions.

Based on information gathered from students, faculty, and Deans, the Task Force has identified the following audiences for end-of-semester course evaluations, and the primary needs of each:

Department Chairs and Deans need timely and reliable information about teaching effectiveness for faculty personnel decisions and mentoring of graduate students. Since they have many faculty to review, they seek information that allows them to quickly compare and evaluate teaching effectiveness; however, this desire may be at odds with best practices for evaluating teaching responsibly and in context. In some Departments and Schools/Colleges, questions from student course feedback surveys may also be used profitably as indirect evidence of student learning in program learning outcomes assessment.⁵

Faculty need timely and reliable formative and summative feedback about their teaching effectiveness in order to improve their teaching and their courses and to better understand student perceptions of their learning. **Faculty who are new to teaching and graduate students** in a teaching role require additional mentoring and formative feedback about their teaching from students, peers, and faculty mentors in order to support their development as teachers.

Students need access to reliable information about courses in order to make informed course selections. This includes timely and reliable feedback about teaching effectiveness and ready access to course syllabi. Students also need adequate understanding of and guidance about the purpose of providing course feedback and how results are used, so that they

⁵ In the College of Engineering and School of Law, this information is systematically gathered from course evaluations for accreditation purposes.

participate by filling out midterm and end-of-semester course feedback surveys responsibly and constructively.

In each of these categories, the need for reliable information is key, and end-of-semester student course feedback alone, particularly in its current form, is insufficient. Over the course of the year, the Task Force consulted a far-ranging body of research on evaluating teaching to conclude that Boston University can and should incorporate a more systematic approach to using multiple measures at different junctures during the semester and over the course of the teaching career. The recent [report](#) of the Association of American Universities, mentioned above, “Aligning Practice to Policies: Changing the Culture to Recognize and Reward Teaching at Research Universities” (AAU 2017), notes that at research universities, teaching effectiveness tends to be assessed primarily based on end-of-semester student course evaluation surveys, despite concerns about what student course evaluations actually measure. Our survey of BU Department Chairs and Deans revealed a similar overreliance on student course evaluations to measure teaching effectiveness, as well as an overreliance on omnibus ratings of instructor effectiveness (e.g., “Overall, how would you rate this instructor” on a 1-5 scale).⁶

Teaching effectiveness cannot be measured or represented as a single number or average score on student surveys, although it can be tempting to reduce it to such metrics. As mentioned above, increasingly, research on the validity of student ratings of instruction and growing concerns over gender and racial bias⁷ suggest that feedback from students should be supplemented with additional measures for evaluating teaching, including a robust and well-defined system of peer observation and peer review of teaching materials and the judicious use of teaching portfolios in all high-stakes personnel decisions. End-of-semester and midterm student course feedback surveys should be viewed primarily as important sources of formative feedback that can be used to improve teaching.

⁶ The literature on global subjective evaluation suggests that when individuals are asked to make global ratings by integrating different pieces of information into an overall score, they tend to do a poor job of integrating information and instead rely on a “gut” response (global sentiment), making these questions more likely to be influenced by characteristics such as likeability, attractiveness, or other physical attributes.

⁷ The Task Force spent considerable time reviewing research studies and recent press on the subject of bias in student ratings of instruction (see Bibliography). There is enough evidence that bias is a concern to warrant devising an approach to evaluating teaching that would reduce its effects, both in the design of a student feedback survey instrument and in considering the place of student feedback in the overall evaluation of teaching, particularly for high-stakes personnel decisions.

The AAU “Aligning Practices to Policies” report, as well as our own research, indicates that many research universities, including Stanford, Yale, Vanderbilt, Brown, UC Berkeley, and University of Southern California, have taken important steps to revise their course feedback forms and reduce reliance on student feedback surveys in the overall evaluation of teaching. Boston University should seize this opportunity to take its place among these leaders.

The Task Force recognizes that shifting emphasis away from a single, numerical metric in the evaluation of teaching would constitute a significant change from the University’s current approach. To support this change, our report includes recommendations about process and, in many cases, sample materials to be used or adapted by Departments, many of which have been developed through AAU research initiatives at member schools. Additionally, Boston University’s Center for Teaching and Learning offers faculty an array of resources, from a mid-semester Learning Analysis Poll, to assistance with selecting, curating, and reflecting on teaching materials for teaching portfolios, to one-on-one consultations to review and interpret the results of student feedback surveys.

Implementing a well-designed process of peer review and review of teaching portfolios for personnel decisions will add to the time faculty and department chairs currently devote to evaluating teaching. The peer evaluation timelines and rubrics and guidelines for teaching portfolios included in this report will help Departments craft an approach, but the University will need to provide guidelines to support Departments and create a reasonable timeline for implementing the new approach. The Task Force believes that such steps are necessary to promote excellent teaching as a core value at Boston University, and to remain aligned with our local and AAU peers.

Key recommendations:

1. Boston University should adopt a university-wide online student feedback survey that is brief and includes both quantitative and qualitative elements. In addition to the core questions given to every School and College, additional, customizable questions can be included at the discretion of individual Schools, Departments, and faculty members.
2. Boston University should adopt online course feedback forms following set guidelines for administering the evaluations to ensure maximum response rates (generally, they should be administered in class during a predetermined two-week window to maintain response rates

comparable to paper evaluations). A move to provide course feedback online ensures that all stakeholders receive results quickly and efficiently.

3. The University should require multiple measures in order to responsibly evaluate teaching, including peer evaluation and teaching portfolios, for all personnel decisions, including contract renewal.
4. Midterm course feedback surveys should be made available to all faculty in all courses, ideally through the same online mechanism as end-of-semester surveys. The emphasis of midterm course evaluation is entirely formative, so questions should be customizable, and results would be viewable only by the faculty member teaching the course. (See Appendix for a sample survey.)
5. Results of end-of-semester feedback surveys should be disclosed to the student body. We recommend that comments in response to a single, open-ended question on the standard course feedback form be shared with students along with a summary of quantitative results. Further, there should be a process by which faculty or staff can redact student comments to remove inappropriate language or content. Students should also have the opportunity to share course information through a student-created web page and have access to published course syllabi to use in considering course selection.
6. Boston University should continue to keep pace with initiatives at AAU peer institutions to better evaluate teaching and emphasize its value in tenure and promotion processes and other faculty reward systems.

No single method can evaluate the quality and effectiveness of faculty teaching; however, by thinking about the purposes of and audiences for teaching evaluations and the strengths and limitations of available tools, we can arrive at a system for evaluating teaching that incorporates multiple measures and focuses on improving teaching at Boston University. Such a system will provide a standard, university-wide framework with enough flexibility to align with the needs of Schools and Colleges, Departments, and individual faculty members.

The following report outlines in more detail the substance of these findings and recommendations, which emerged not only from the Task Force's own research, but also from consultations with many stakeholders at Boston University, including Student Government, Undergraduate Student Advisory Board, Graduate Student Advisory Board, Faculty Council,

ARROWS, Undergraduate Council, Graduate Council, and Council of Deans, and interested faculty.

Revised Boston University Course Feedback Survey

Following its charge, the Task Force has drafted standard content for a university-wide course feedback form to replace the widely divergent instruments currently used in the different Schools and Colleges on the Charles River and Medical Campuses. Although we recommend that scores on student course feedback surveys be viewed as only one measure among many to evaluate teaching, we continue to see these surveys as an important source of feedback for improving teaching quality, especially if Schools and Colleges, Departments, and individual faculty members retain the ability to pose questions that are highly-tailored to their own needs and pedagogical approaches.

In its 2009 report, the Hanover Research Council examined the advantages and disadvantages of standardized and non-standardized course evaluations. The majority of institutions in its study of *US News* top-ranking national universities and liberal arts colleges use a “mixed” format, where some questions were standard across the university and others were tailored to the needs of departments and faculty members. Standard forms are advantageous because they provide information that can be compared across departments, schools and colleges as well as standards that are consistent from year to year. UW Madison reported that using a standard form protects untenured faculty members from inconsistencies or biases within individual departments (Hanover 3). However, non-standardized forms are also advantageous because they allow for cross-disciplinary variance and varying departmental needs. Additionally, using different formats may help better engage students in the serious task of providing responsible feedback. As a compromise between institutional needs and the potential needs of departments and individual faculty members, a mixed approach allows universities to design an instrument to meet the needs of both groups. Many of BU’s AAU peers use a mixed approach, including Northwestern, UC Irvine, Yale, and UC Berkeley.

In light of current debates about gender and racial bias in student course evaluations and some research findings about the lack of correlation between scores on these instruments and student learning (measured by exam scores and other direct measures as well as student success in courses they take subsequently), we approached the task with some misgivings about what quantitative scores on student feedback forms actually represent. As a result, we decided to draft an instrument that would not include questions asking students to rate the instructor personally or judge his or her performance. Instead, we thought it was important that

the instrument **emphasize student perceptions of their own learning** and **focus on how well courses were fulfilling stated learning objectives**. We also wanted to include questions that specifically direct students to reflect on their own contributions to the course as partners in the learning process.⁸

Another consideration, based on research findings at other institutions, was to design a form that would not encourage “straight line” responses from students. A recent study of student responses in end-of-semester course feedback surveys at Stanford University indicated that 35% percent of students routinely assign the same instructor rating in response to all questions on the survey, following a straight line down the page (Stanford 2013). **Our proposed course feedback form does not lend itself to such “straight line” responses**, since each question or group of questions requires a different type of response, and not all use the same scale.

The same study found that although students were being asked to rate up to 38 discrete course elements, responses broke down along only four dimensions, suggesting that the feedback form could be drastically shortened without compromising the results (Stanford 2013). At Boston University, a simple coding of current question types revealed some redundancy on our current survey instruments. Based on these and other similar findings elsewhere, **our proposed course feedback form is also shorter than most of the instruments currently being used at BU**.

The results of our inventory of current BU question types showed that most departments rely primarily on quantitative rating scales with a small number of free-response questions. Based on our inventory, two of the most frequently-asked questions on student course feedback surveys at Boston University are “How would you rate this course overall?” and “How would you rate this instructor overall?” Another frequently-asked question invites students to rate how well the course prepared them to meet specific learning objectives at either the course or the program level, although there was variance in the degree of granularity of these learning objectives (ranging from “Please rate the extent of your learning” or “Rate the extent to which course objectives were achieved” to “How well did this course prepare you to write clear, correct prose?” or other specific, accreditation-driven program outcomes). Other high-frequency

⁸ A research study at conducted at the University of Rochester revealed significant bias in student responses correlating with expected grade, prior interest in the course, and instructor attractiveness (Rogge and Zahn, “University of Rochester Course Evaluation Project” 2011).

questions ask students to rate an instructor's knowledge, enthusiasm, and ability to motivate students to participate in class. The most frequent open-ended, and possibly most useful response questions are "What were the strong/weak points of the course?" and "What suggestions do you have to improve this course?"

Many, but not all, departments currently find it useful to ask questions relating to course materials, classroom (or online delivery) technology, assignments, particular instructional modules, guest lecturers, and other instructional components. In our revised course feedback survey instrument, we recommend that Schools, Programs, Departments and individual faculty members continue to ask these questions using the customizable section of the course feedback form, as these questions are most likely to elicit concrete suggestions to be used to improve courses and teaching. (This level of customization will require departments to supply information about their courses, such as learning objectives and other features to be evaluated, in advance, so that the information can be included in the course feedback form students receive. A process for this will need to be devised, as it has been at our peer institutions using a mixed approach to course evaluation).

Our inventory revealed that most departments do not currently ask questions relating to student effort and engagement in the course. These include questions that ask students to estimate the time and effort they put into the course, estimate their expected grades, or otherwise evaluate their own performance. Some ask students for demographic information that could be captured through an online course feedback system. We believe it is important to include questions about effort and engagement in the proposed student feedback survey so that students reflect on their own responsibilities for classroom learning, and have added a section called "student effort" in the new instrument.

In drafting a university-wide student feedback form, we consulted numerous models and examples from other universities and reviewed literature, including national inventories of question frequency on university-level SRIs (Student Ratings of Instruction) and the for-fee IDEA Center Student Ratings of Instruction System. We were particularly drawn to approaches that included a mix of quantitative and open-ended responses and did not lend themselves to oversimplified numerical reduction. Another important consideration revolved around the questions themselves. Feedback surveys should include questions about only those course elements that students are qualified to evaluate. In order to reduce bias, questions should emphasize student learning over instructor attributes. As mentioned above, we also feel

strongly that the feedback form should include questions that emphasize a student's own accountability for learning by encouraging them to reflect on their own level of effort.

With those considerations in mind, we have retained some questions on the standard course feedback form that were high-frequency questions based on our inventory of current course surveys on campus, while maintaining a focus on student learning and on elements of a course that students are best-equipped to evaluate. Because our survey revealed that BU departments rely heavily on so-called “umbrella” questions, we have retained one question that asks students to rate the quality of instruction overall, despite some concerns that such questions are most susceptible to misinterpretation and/or bias (they are sometimes called “popularity contest” questions). We hope that by asking students to rate the quality of instruction and not the instructor, and then asking students to rate the course overall in the context of similar courses at the same level or of the same type, the re-worded questions will encourage more thoughtful, nuanced responses. ***We would like to emphasize that these questions should not be taken out of the context of the rest of the survey as stand-alone, quantitative instructor “ratings.”***

In interpreting results of student feedback surveys, no one numerical score should be taken out of context as a “global” score to numerically represent a faculty member's teaching effectiveness, as the questions are designed to work together to provide a contextualized view of students' perceptions for the primary purpose of giving feedback to improve teaching.

Finally, although we have relied on questions and rating scales that have been used and tested at other universities, we recommend piloting and testing the new question set in order to assure its validity and reliability. The ongoing SPH pilot will provide useful information.

We envision that the customizable section of the feedback form be used primarily to ask questions about particular learning goals and instructional elements, with the caveat that departments add these questions judiciously and consider the overall length of the resulting survey. A bank of recommended questions to select among will be provided. The customizable section of the feedback form could also include a standard module to be used for BU Hub courses approved for each Hub Area to evaluate how well students feel a course prepared them to meet the learning outcomes set out for them (see example below). Feedback from students on these questions can be used both by faculty and departments and also more broadly for assessing student learning in the BU Hub university-wide.

Proposed University-Wide Student Course Feedback Form

Questions common to all feedback forms:

Course Content and Instruction:

1. How organized did the course seem to you? (*extremely organized, very organized, moderately organized, slightly organized, not organized at all*)
2. How well were the course learning objectives communicated? (*extremely well, very well, moderately well, slightly well, not at all*)
3. How well did the course fulfill its stated learning objectives? (substitute up to 4 specific learning objectives) (*extremely well, very well, moderately well, slightly well, not at all*)
4. How much did you learn from this course? (*a great deal, a moderate amount, some, a little, nothing*)
 - 4a. Please explain your response. (free response)
5. What were the most valuable aspects of the course? (free response)

Student effort:

6. What percentage of class meetings did you attend? (0-20, 20-40, 40-60, 60- 80, 80-100)
7. Outside of class, how many hours per week on average did you spend preparing for class? (0-4, 4-8, 8-12, 12-15, 16+)

Student learning

8. What skills or knowledge did you learn or improve in this course? (open-ended)

For students:

9. What advice would you give to students considering taking this course in the future? (free response)

Additional Feedback:

10. If you have additional comments about this course, please provide them below.

Customizable questions:

Additional questions may be added to the core feedback form to meet the needs of Schools and Colleges, Departments, and individual instructors. These should be added judiciously, with the goal of collecting feedback that will improve teaching. Some recommended questions include:

How well did you achieve this learning goal in the course (instructor may add up to 10 learning goals; if none are added, a generic question will be asked: "How well did this course help you meet the stated learning objectives")

How useful to you was this course element? (department or instructor may identify up to 8 elements to evaluate, such as lectures or guest lectures, readings, problem sets, group projects, textbook, online component, technology tool, or other)

Optional additional faculty-posed questions (up to 3 additional of each, open or closed response)

**Sample BU Hub Course Evaluation for a course approved for
Philosophical Inquiry and Life's Meanings and Historical Consciousness**

In addition to questions 1-11,

1. How well did this course develop your knowledge of notable works in philosophical thought?
2. How well did this course develop your ability to make meaningful connections among philosophical works?
3. How well did this course develop your ability to relate philosophical works to your own life and to the lives of others?
4. How well did this course help you develop the vocabulary and reasoning skills to reflect upon significant philosophical questions and topics?
5. How well did this course help you understand the characteristics of at least one literary or artistic medium and produce evaluative, analytical, or creative works to demonstrate that understanding?
6. How well did this course help you learn to create historical narratives, evaluate interpretations based on historical evidence, and construct historical arguments?

The University should develop a set of standard questions for BU Hub courses based on the learning outcomes for those courses. Similarly, courses with lab components would add a section asking students to give feedback on that component of the course. We noted that some universities maintain a general question bank of possible questions for faculty to include in the customizable section of the survey instrument, and BU could consider developing such a resource. Departments should not use the customizable section of the feedback form to simply reintroduce their current course evaluation questions.

Collecting Course Feedback Online

The Task Force recommends that BU adopt an online system of gathering student feedback on courses. Many departments at Boston University are still using scanned paper evaluations to collect student feedback, in part because faculty have been concerned about moving to an online course survey system.

It is important to consider their primary concerns in turn. First, research has shown that student response rates often go down when universities move to online systems, unless they make a concerted effort to administer evaluations in class during a regularly scheduled class meeting. If this step is taken, there is no significant difference in response rate between the two methods. Faculty also worry that online feedback surveys encourage students to see themselves as consumers, decreasing the likelihood that they will understand the importance of providing thoughtful, responsible feedback about their courses. Faculty members fear that both the quality of response) and the quantity of open-ended feedback will go down. Although the concern is that students may not write as much when they have to type on a keyboard or hand-held device, research has shown that there is no significant difference in the amount of feedback that students provide online (Chang 2003 and Norris 2005). Students with higher GPAs and female students are somewhat more likely to respond to online evaluations, and open-ended comments are actually significantly longer in online course evaluations, providing meaningful feedback for all stakeholders (Crews and Curtis 2011 and Collings 2004, "Online student survey comments: A qualitative improvement?"). A 2002 study showed that students were more likely to respond when they believed ratings would be used in decision-making (Johnson 2002). This finding underscores the importance of communicating to students how information from course feedback surveys will be used to improve teaching, ideally on the survey itself, and also to give them access to results. At the institutions where these studies were conducted, students, faculty, and administrators all expressed satisfaction with online course feedback surveys.

Perhaps the most important advantage of online course feedback surveys is the speed with which feedback can be provided to faculty, making it possible to incorporate feedback quickly into current and upcoming courses. Online tools also provide better data and reporting to perform statistical analyses with less manipulation error and provide universal access to all students. Lastly, online course feedback surveys are less expensive and more environmentally

sustainable compared with the processes and materials associated with paper evaluations. Nearly all of BU's AAU peers have made the move to online course feedback surveys (see attached table of peer comparisons), and the Task Force recommends that Boston University do the same, taking care to put in place several measures to maintain high student response rate:

1. Administer course feedback surveys in class, during a regularly scheduled class meeting.
2. Leave the window for evaluation open for two weeks so that students can continue to provide feedback outside of the classroom and to accommodate all students.
3. Provide students with standard framing language describing how feedback from the surveys will be used and highlighting their role and responsibility in the process.
4. Consider providing incentives to students who complete online course evaluations, such as early access to grades or to results. Some universities only allow students who have submitted course feedback to view results; BU could consider such an approach.

During the 2017-2018 academic year, the Task Force monitored a pilot implementation of a vendor-based (Campus Labs) online course evaluation system in the College of Arts and Sciences, College of Communication, College of Fine Arts, Law School, Metropolitan College, and Wheelock School of Education. Following the Fall 2017 pilot, an analysis of all participating schools showed an overall response rate of 80%. Of 2884 CAS and GRS course sections, there was a response rate of 71%. There was 87% response rate in Wheelock, 73% response rate in COM, and a 90% response rate in the School of Law. A faculty feedback survey and comments from a focus group of academic administrators indicated general satisfaction with the move to an online process. Faculty particularly appreciated receiving results quickly and the reporting features of the Campus Labs system. Some initial concerns about lower scores on online surveys compared with paper ones do not seem to be borne out, upon further investigation. From the feedback we received, it seems clear that faculty are satisfied with online course evaluations in general and recognize the advantages, but that aspects of the pilot implementation with the Campus Labs system were less successful due to issues with properly identifying courses, challenges with cross-listed courses, and technical issues with the platform itself during the evaluation period. A uniform set of guidelines and procedures and a system that is fully integrated with the student information system (which was not in scope for the pilots) should resolve many, if not all, of the issues reported.

Midterm Course Evaluations

To complement the revised end-of-semester course feedback survey, we recommend that all instructors collect midterm course feedback, and that a survey mechanism be made available to all instructors using the same online course platform used for the end-of-semester surveys. Overall, traditional end-of-semester course evaluations arrive too late for students enrolled in a particular class to benefit from the feedback they provide on the evaluation and are typically standardized across the college/university to allow for cross-discipline comparisons (Ragupathi, n.d.). The midterm survey is a valuable opportunity for faculty to gather customized feedback while they are still able to make adjustments to the course.

Mid-semester evaluations allow students to share their experiences of a course typically in weeks 6-8 of a 15-week semester and also allow instructors to modify the course as needed to meet the needs of learners. Mid-semester evaluations provide ongoing feedback, which allows instructors to modify a course's organization, experiment with instructional strategies, and get concrete feedback about student learning experiences. The findings from a mid-semester process are confidential and are used constructively by the instructor. Additionally, they "provided insights into students' perceptions of how learning works that are not identified in the end-of-course student evaluations of instruction" (Dangel & Lindsay 27). Mid-semester evaluations are flexible; they can be administered by instructors or externally by a Center for Teaching & Learning. The instrument can be adapted to gather qualitative or quantitative feedback about a specific class, course, or module. While the instrumentation is flexible (in principle, it may be paper-based in class, e-mail facilitated, or administered through an online form), delivery through the same online platform used for the end-of-semester survey would streamline the process and ensure that it is centrally available to all instructors and therefore encourage adoption.

Research indicates that instructors using mid-semester evaluations improved their teaching. In a 2004 study the researchers found that instructors, "clarified expectations of students, and refocused content emphasis. They also indicated that they intended to amend the way they teach future courses in an effort to increase effectiveness" (Diamond 217). In a 2011 study at Brigham Young University, researchers completed a systematic analysis of 249 mid-semester evaluations including 305 faculty and 3,550 students to understand the perceptions of the effects of mid-semester evaluations on teaching and learning. The results indicated that both faculty and students perceive improvements in student learning and faculty teaching and

that “small changes in teaching may lead to large improvements in student perceptions of their learning” (McGowan & Osguthorpe 170).

To supplement the proposed online mid-semester survey tool, Boston University’s Center for Teaching & Learning (CTL) offers a mid-semester evaluation process for instructors called the Learning Analysis Poll (LAP). The qualitative instrument is administered by a CTL staff member during 30 minutes of a standard class period. Student groups generate feedback about what they perceive is working in the classroom and what they perceive is not working in the classroom, and they offer recommendations for improvement. The data is then shared at the class level for consensus. Only responses that have been agreed upon by a majority vote are reported, so the instructor knows that most students concur with the suggestions. The data is further analyzed by CTL to provide evidence-based recommendations to instructors.

Balancing Student Course Feedback with Other Evidence of Teaching

“There is no known objective method for measuring teaching performance. The trick is to find a sound, defensible strategy for combining multiple sources of 'informed' judgment for reaching those decisions.” - Ronald Berk, Thirteen Strategies to Measure College Teaching

Responding to criticism that student evaluations of teaching (SETs) are often biased, statistically unreliable, and the sole teaching-related data point that factors into faculty retention and promotion decisions (American Association for the Advancement of Science, 2012; Berk, 2006; IDEA, 2011, Association of American Universities “Aligning Practice to Policies” 2017), many institutions have sought to develop more holistic and reliable methods for measuring teaching effectiveness. Despite their drawbacks, however, almost all agree that end-of-semester course feedback is imperative to any well-balanced evaluation. What other forms of evidence, then, are necessary to conduct a fair and *useful* evaluation of teaching effectiveness? Further, how much should these additional data points count in important decisions about contract renewal, merit, retention, and promotion?

The criteria by which teaching quality is measured need to capture teaching practices and behaviors such as instructional effort, student engagement, and design for learning (Balanced Evaluation of Teaching Model, n.d.; Berk, 2006) - data that traditional student feedback surveys rarely collect. One argument behind this assertion is that student feedback surveys provide little information about the specific teaching and learning context, which can heavily influence teaching decisions and learning experiences. The omission of contextual data makes evaluating the objectivity of student responses difficult at best. Attempts by the instructor to integrate new or evidence-based teaching models may stumble at first, for example, or learner-centered classroom activities may unsettle students who have performed well in traditional learning environments. In both instances, an instructor may receive lower ratings from students, despite efforts to improve teaching practice, thereby discouraging further attempts at pedagogical change. Our aim is for the customizable elements of the proposed university-wide student course feedback survey to allow faculty to collect such contextual information and feedback relevant to their own pedagogies, goals, and experiments.

Other evidence of teaching effectiveness to be considered in the overall evaluation of teaching include peer review of teaching materials, peer observation, and a faculty member’s own reflection on teaching effectiveness supported by a teaching portfolio and statement of

teaching philosophy. It will be important to engage *faculty themselves* in developing the specific instruments and procedures for measuring teaching effectiveness (Vasey & Carroll, 2016) at the department level to reflect department standards and disciplinary norms. Not only does participation in the instrument development process encourage ownership and greater buy in on the part of instructors, it opens what is typically a black box of the evaluation process for critique and refinement.

The recent findings of the AAU recognize that “a central challenge to enabling effective evaluation of teaching practices is provide a practical framework that is scholarly, accessible, efficient, and aligned with local cultures” (AAU “Aligning Practices to Policies” 2017 and Wieman, “A Better Way to Evaluate Undergraduate Teaching,” *Change* 2015). The sections that follow elaborate on the essential components of a holistic evaluation process for departments to consider. More specific tools and resources (such as rubrics and peer observation guidelines) appear in the appendices.

Teaching Portfolios

Teaching portfolios bring together a variety of materials documenting an instructor’s teaching effectiveness. Portfolios typically begin with a brief reflective statement that outlines the instructor’s pedagogical goals or approach to teaching, and include annotations that contextualize the accompanying materials, which provide “evidence” for the claims made in the reflection. Teaching portfolios are meant to be representative, not exhaustive, and context will dictate which materials should be included. Materials might include:

- Overall reflective statement
- Statement of teaching philosophy
- Description of development of and changes to pedagogy and teaching environment (for example, moving to a studio classroom or shifting to an online course)
- Commentary on student feedback surveys and any changes made as a result
- Sample syllabi and course materials
- List of courses taught, teaching workshops or training attended
- Evidence and examples of student learning
- Discussion of contribution to department curriculum development
- Scholarly work on teaching and learning

Teaching philosophy statements, which are increasingly requested by committees considering a candidate for tenure and promotion, typically explain an instructor’s learning goals for students; the educational strategies used to help students reach those goals; and assessment methods by which the instructor measures student progress towards those goals.

Anchored in specific personal examples, teaching statements help readers to understand how an instructor teaches, and why the instructor teaches that way.

CTL consultants can work one-on-one with faculty to develop and refine a statement of teaching philosophy. The process of writing a teaching statement provides instructors with an opportunity to clarify their classroom goals, such as creating an inclusive classroom or educating students about the responsible use of technology; to evaluate whether their teaching practice aligns with their goals; and to reflect on their pedagogical development.

Reflective annotations allow instructors to offer their reader a fuller picture of their teaching than quantitative data alone could provide. For example, an instructor might reflect on what she learned from teaching a course that produced low student feedback scores and note any subsequent actions that she took (consulting the CTL about how she could adjust her teaching to better align the expectations of this student body, observing peers teach, etc.).

Teaching portfolios can be paper-based, but increasingly, scholars are showcasing their teaching experience in online portfolios. The Educational Technology team can help faculty to create their digital teaching portfolio using the Digication platform, and there are other free online portfolio platforms, including Wix, Weebly, and Wordpress. A practical and widely-used guide to teaching portfolios, now in its fourth edition, is Peter Seldin's *The Teaching Portfolio* (2010). CTL staff can help instructors one-on-one with the selection and curation of materials and the process of reflecting on these materials.

The Task Force recommends that Departments consider materials put forward by faculty members themselves in a teaching portfolio as an important part of a holistic evaluation of teaching.

Peer Review

Peer observation and peer review of teaching materials is an essential part of the proposed holistic framework for evaluating teaching. Observing fellow teachers can both magnify the impact of student feedback as well as more directly measure teaching effectiveness (Cohen, 1980; Menges & Brinko, 1986). Peer observations can reveal environmental factors not always present in student evaluations, including the presence or absence of a centralizing classroom climate, active learning, and instructional scaffolding (Ambrose et al., 2010). Nancy Chism's

Peer Review of Teaching: A Sourcebook, is a well-regarded resource for departments to consult. For observations to be effective, they typically involve the following steps:

- Preparatory meeting (prior to the classroom visit)
The observer collects information about the course context.
Review syllabus.
Discuss the types of learners in the class.
Discuss format and methods of instruction for the class.
Discuss how feedback is provided to students.
Discuss areas of focus for the evaluation (share rubric).
Other areas, as requested by the faculty member being evaluated.
- Observation:
The observer completes a standard peer evaluation form.
Faculty member being observed completes a self-appraisal to be used as part of the post-evaluation meeting .
- Follow-Up meeting Analysis:
Peer observer and faculty member meet to discuss the evaluation and self-appraisal.
A summary should be jointly developed that may include strategies for improvement, as appropriate (Brinko, 1991; 1993).

Many who responded to the Task Force's survey of Boston University Deans and Department Chairs indicated that they would like to have better tools for peer evaluation of teaching. A related finding was that a number of departments and Schools have guidelines for peer observation that are not being regularly put into practice. These findings mirror national findings about the desirability and implementation of peer observation to evaluate teaching. Researchers at Cornell and Yale conducted an Effective Evaluation of Teaching and Learning Survey, which asked respondents to describe current methods being used to evaluate teaching and learning at the undergraduate level and perceptions of the desirability of several best practices for assessing teaching effectiveness.⁹ Not surprisingly, 95% of respondents, which were faculty and department chairs in STEM disciplines, indicated that their institutions measure teaching quality with paper or online student evaluations. Interestingly, when those same institutions were asked what teaching evaluation methods would be desirable, greater than 95% percent felt that peer observation of teaching would be highly or moderately desirably, even at

⁹ The survey received 324 responses from STEM department faculty drawn from both primarily undergraduate institutions (67%) and research universities (33%), including 38 department chairs (AAU Searching for Better Approaches 2015). The other measure that 91 % the group found highly or moderately desirable for evaluating teaching effectiveness is longitudinal assessment of student learning, but only 5% reported "always" or "regularly" implementing this approach.

research universities, whereas only 50% reported it was used “always” or “regularly” in their departments (AAU 2015, 15), revealing a striking implementation gap between evaluation practices perceived to be desirable and those actually used. This implementation gap represents an excellent opportunity for institutional change at research universities.

The primary concerns from faculty about regular peer observation relate to the time it takes, a perceived lack of expertise on the part of observers, and lack of established criteria for evaluation or consensus about what constitutes good teaching. We hope that providing some sample materials will help mitigate at least some of these issues.

Two models we encountered bear mentioning in more depth, as they provide a concrete framework for implementing peer review of teaching to help contextualize feedback received from students. The first is a recommended [protocol](#) from NC State University (see Brent and Felder, “A Protocol for Peer Review of Teaching” *Proceedings for the 2004 American Society for Engineering Education*), based on the results of a pilot peer observation program that was subsequently adopted at the university:

- 1) Design class observation and course material rating forms (examples in Appendix E-Appendix H) and obtain consensus approval of the department faculty for the items included in the final forms
- 2) At the beginning of the fall semester, form a departmental peer review committee that will function for the next academic year. The committee should consist of a chair within the department who oversees the process and a cadre of faculty raters who may come from within the department or from other departments in related disciplines.
- 3) Early in the fall, provide a 1-2 hour training session to the raters using an illustrative set of course materials and one or two mini-lectures or videotaped excerpts of classroom teaching.
- 4) Summative review. For faculty members being considered for reappointment, promotion, or tenure, two raters should visit the same two classes, reconcile their scores, review teaching materials, and create a summary for the chair of the peer review committee.
- 5) Formative review. Observations can be conducted by only one rater who shares and discusses the results only with the instructor. Such constructive feedback provided

to faculty members in their first few years of teaching should help them meet or exceed standards for teaching in their departments.¹⁰

Notably, the total time required for a summative review using this suggested protocol is approximately seven hours per rater, similar to a typical committee assignment.

The second model comes from the University of Oregon, where holistic evaluation of teaching is now mandatory. Their “Proposed UO Peer Review of Teaching Framework” from April 2018 is included in Appendix H. As in the NC State framework, each department identifies a peer review coordinator to oversee the process. At UO, temporary faculty would be reviewed once per year, career instructional faculty are reviewed once per contract period, assistant professors are reviewed once per year, associate professors are reviewed once every other year, and full professors are reviewed once every three years. The UO Provost’s [website](#) outlines these requirements for all faculty. The Task Force agreed that a schedule such as this would make sense at Boston University, but would recommend allowing departments to opt in at first. As at NC state, UO asks each department to either create a peer observation tool or use a customizable tool provided by the university. Each department completes a brief peer review report for each faculty member describing information collected from classroom observation, faculty self-assessment, and summing up discussion at a follow-up meeting. For BU, the Task Force recommends also including a review of teaching materials to the scope of review and reporting.

Through our research, we noted that universities consider peer review and observation to be the ideal way to evaluate teaching, but that many struggle to implement it systematically. In our view, the best approach is to provide a simple framework and a set of materials that departments can easily customize as needed to reflect their own needs and disciplinary standards. In addition to the documents pertaining to the specific practices at NC State and University of Oregon, we are also including two relevant sample documents referenced in the 2017 AAU report. These are a peer evaluation of [teaching rubric from the University of Kansas](#) (Appendix E) and a copy of UC Boulder’s [Teaching Quality Framework](#) (Appendix F), which incorporates evidence from student feedback surveys, faculty peer review, and individual instructor reflections (such as a teaching portfolio).

¹⁰ For further details on this protocol, see Rebecca Brent and Richard Felder, “A Protocol for Peer Review of Teaching” in *Proceedings of the 2004 American Society for Engineering Education*, 2004.

Conclusion

The Task Force on Evaluating Teaching's new, university-wide course feedback form with both standard content and flexible options for Schools/Colleges, Departments, and individual faculty members to be delivered online, with results made available to students, will put Boston University in line with excellent teaching evaluation practices nationally. The new survey tool, when viewed alongside other measures of teaching effectiveness such as systematic peer observation and peer review of teaching materials and feedback/reflection from individual instructors, will create a framework for evaluating teaching that is both rigorous and flexible enough to reflect disciplinary and departmental culture. Adopting these measures will put BU at the forefront with AAU peers making similar adjustments to how they value and evaluate teaching institution-wide.

Our suggested framework for evaluating teaching provides structure, while allowing for local variation in a manner that preserves academic freedom for faculty and recognizes differences in disciplinary and departmental culture. Our proposal is intended to provide high-quality evidence of and feedback about teaching effectiveness to all stakeholders to help foster a university culture that values excellent teaching and continuous improvement of teaching.

Appendix A: Audiences for Evidence of Teaching Effectiveness

Stakeholder	Information/Evidence Needed	Resources Needed to Support the Evaluation Process
Department Chairs and Deans	Results of end-of-semester course evaluations, results of peer review of teaching, faculty teaching portfolios, and in some cases, aggregated results of questions related to program learning outcomes	<p>Clear examples and models to guide the process of peer review and classroom observation</p> <p>Suggested criteria for faculty teaching portfolios</p> <p>Assistance drafting customized questions at the Department or School/College level</p>
More Experienced Faculty	Results of midterm course evaluations, end-of-semester course evaluations, peer review of teaching materials and classroom observation	<p>Confidential input from a department mentor</p> <p>Assistance interpreting course evaluation results from the Department and Center for Teaching and Learning</p> <p>Guidelines and criteria for assembling a teaching portfolio</p>
Less Experienced Faculty	More focus on confidential midterm and peer review or mentor review	same
Students	Access to results of Boston University's end-of-semester course evaluations to aid in course selection, a syllabus repository, evidence that results of midterm and end-of-semester course evaluations are being used to improve teaching.	<p>Information about the purpose and audiences for course evaluation</p> <p>Guidelines for how and why to participate in course evaluation responsibly</p>
Graduate Students	Same as above for both faculty and students.	<p>Active mentoring from Departments, from interpreting the results of course evaluation to feedback on classroom performance and assistance compiling teaching portfolios.</p> <p>Outside of their departments, opportunities for professional development tailored to their needs from CTL and Office of Postdoctoral Affairs.</p>

Appendix B

Schedule of Consultations

June 23, 2017--Student Government

September 8, 2017—Student Government

September 19, 2017 – Undergraduate Council

November 9, 2017 – Graduate Council

December 15, 2017 – Undergraduate Student Advisory Board

January 17, 2018 – Council of Deans

March 13, 2018 – Faculty Council

March 14, 2018 – Graduate Student Advisory Board

April 4, 2018—BU ARROWS

August 23, 2018—Associate Provost for Diversity and Inclusion

Appendix C

Course Evaluation Practices at Selected Peer Institutions

Institution	Online	Shares Results with Students/University Community	Incentives for Students	Redacts results	Process includes Peer Evaluation	Standard survey	Outcomes-focused course survey
Northeastern	Yes	Yes	No	Yes		Yes	
Boston College	Yes	Yes—summary responses to Part A of the survey (no free response)	No	No (free responses not shared)		Yes	
Brandeis	Yes	Yes, through the “Brandeis University Course Evaluation Guide” selection portal	No			Yes	
MIT	Yes	Yes	No			Three different instruments (by discipline)	No
University of Colorado Boulder	Yes	Yes	No		TQF (Teaching Quality Framework) initiative (NSF and AAU funded)	yes	No
Duke University	Yes	Yes, numerical results only	No	Faculty may choose not to opt in to sharing results			Yes
Stanford University	Yes	Yes	Yes. Students who participate receive early		Yes	Yes, with flexible sections (mixed format)	Yes

			access to grades				
University of Kansas	Yes	No	No	N/A	Yes—NSF funded project for holistic evaluation	No	variable
UMass Amherst	Yes	No	No	N/A	Yes. Recommend multiple measures to evaluate teaching	Yes	No
University of Oregon	Yes	Yes	Yes, early access to grades		Yes-peer review required according to a set schedule	Yes; Instrument currently under revision	
University of Southern California	Yes	No—these are used solely by faculty to improve teaching	N/A	N/A	Yes; moving to a holistic approach (student evaluations no longer to be used in tenure and promotion process, citing concerns over bias)		Yes—newly revised survey focuses on student learning
New York University	Yes	Yes				Yes	
Northwestern	Yes	Yes	Yes-only students who respond are able to view results		calls for reform of the system to move to a more holistic process using multiple measures		
University of Michigan	yes	selected results shared via an interactive tool to aid in course selection	No, but faculty are encouraged to give bonus points or other			Yes- 8 university-wide “core questions” and bank of allowable	

			small incentives for participation at the course level—for example, everyone in the class receives a bonus point if the class reaches a 90% response rate			additional questions (phasing out “overall, this was an excellent course” and “overall, this was an excellent instructor”), plus up to 5 instructor-composed questions	
Vanderbilt U	Yes	Yes, no free responses	No	N/A	Only summative are required; formative are encouraged		
Yale University	Yes	Yes (questions 1-9 only)	No	No		Yes. Instructor-specific questions are handled using a paper process.	No

Northeastern: <https://www.northeastern.edu/trace/letter-from-the-2011-2012-faculty-senate-trace-implementation-committee/>

Boston College: https://www.bc.edu/offices/stserv/academic/online_course_evals.html

Brandeis: <http://www.brandeis.edu/provost/faculty-info/courseevaluations/>

MIT: <https://registrar.mit.edu/classes-grades-evaluations/subject-evaluation>

UC Boulder: <https://www.colorado.edu/fcq/>

UC Boulder: <https://www.colorado.edu/teaching-quality-framework/about-tqf>

UC Boulder: <https://www.colorado.edu/today/2017/09/29/cu-boulder-leads-national-study-promoting-next-gen-teaching-evaluations>

Duke: <https://assessment.trinity.duke.edu/students-course-evaluations>

Duke: <https://learninginnovation.duke.edu/faculty-opportunities/art-and-science-of-teaching/assessing-your-teaching/>
Stanford: <https://registrar.stanford.edu/students/online-course-evaluations>
Stanford: <https://evals.stanford.edu/end-term-feedback/course-feedback-form>
Kansas: [https://cte.ku.edu/sites/cte.ku.edu/files/docs/KU Rubric for Evaluating Teaching DEC2016.pdf](https://cte.ku.edu/sites/cte.ku.edu/files/docs/KU_Rubric_for_Evaluating_Teaching_DEC2016.pdf)
UMass Amherst: <https://www.umass.edu/oapa/student-response-instruction-srti>
UMass Amherst: <https://www.umass.edu/oapa/srti/srti-and-performance-appraisal>
Oregon: <https://provost.uoregon.edu/peer-review-and-evaluation-teaching>
University of Southern California: <http://cet.usc.edu/resources/instructor-course-evaluation/>
University of Southern California: <https://faculty.usc.edu/>
NYU: <https://www.nyu.edu/students/student-information-and-resources/registration-records-and-graduation/registration/course-evaluation.html>
Northwestern: <https://www.northwestern.edu/ctec/>
University of Michigan: crit.umich.edu
University of Michigan: <https://ro.umich.edu/faculty-staff/teaching-evaluations/whats-new>
Yale <https://registrar.yale.edu/students/course-evaluations/yale-college-student-oce->

Appendix D

Mid-Semester Feedback for Faculty: Three Questions for Feedback

The objective of mid-semester feedback is to offer constructive information to your instructor BEFORE the end of the semester. Only the instructor will see this feedback. Your answers and comments will remain anonymous.

1. What do we do in this course that is most helpful to your learning?

2. What do we do that is least helpful to your learning?

3. What suggestions do you have for change?

Thank you for providing feedback that will be used to improve this course!

Department Evaluation of Faculty Teaching Rubric

This rubric is intended as a guide to providing a richer, more complete evaluation of teaching. It is designed to structure department evaluation of faculty members' teaching, including contributions to individual courses and to the department's curriculum. The goal is to help departments integrate information from three sources:

- Ø **The faculty member** (including course materials, evidence of student learning, and reflections on student learning)
- Ø **Peers** (including class visits, observations from team teaching, evaluations of course materials, and formal discussions about the faculty member's approach to teaching)
- Ø **Students** (primarily course evaluations)

The rubric can easily be adapted by departments to fit particular disciplinary expectations and to weight areas in ways most meaningful to the discipline. It is intended to guide evaluation of faculty members in the promotion and tenure process.

When completing the rubric, evaluators should consider several factors, including:

- **Types of courses taught** (required or elective, major or non-major, lecture or discussion, team taught or individual, size and level of class).
- **Stage of the faculty member's career** (tenure track, tenured, instructor, adjunct)

Departments may focus on various facets of the rubric at various stages in a faculty member's career, but at all times, evidence of student learning should be paramount.

How to use this rubric:

1. After department members review the form, what modifications do they suggest to make it appropriate for their department?
2. What weights should be assigned to each category?
3. Has the department come to a consensus about the questions and criteria outlined in each category? (e.g., in some departments, advising responsibilities have been under articulated, so faculty members should clarify expectations for their department)
4. Have sources of evidence been identified for each category? This should include a framework for how to read student evaluations of teaching and where they will be used as evidence within the rubric.

University of Kansas Rubric for Department Evaluation of Faculty Teaching (departments may modify as needed)

	Below Expectations: 1-2	Meets Expectations: 3	Exceeds Expectations: 4-5
Expectation levels align with KU's promotion and tenure rating scale.	(1) Poor: Consistently at this level (2) Marginal: Some teaching at this level	(3) Competent	(4) Professional: Some teaching at this level (5) Advanced: Consistently at this level
Goals, content, and alignment <ul style="list-style-type: none"> • What are students expected to learn from the courses taught? • Are course goals appropriately challenging? • Is content aligned with the curriculum? 	<ul style="list-style-type: none"> • Course goals are unclear, inappropriate, or marginally related to curriculum • Content and materials are outdated or unsuitable for students in the courses • Range of topics is too narrow or too broad • Content is not clearly aligned with curriculum or institutional expectations 	<ul style="list-style-type: none"> • Course goals are articulated and appropriate for curriculum • Content is current and appropriate for topic, students, and curriculum • Course topics include an appropriate range • Standard, intellectually sound materials 	<ul style="list-style-type: none"> • Course goals are well articulated, high quality, and clearly connected to program or curricular goals • Content is challenging and innovative or related to current issues and developments in field • Topics are of appropriate range and depth, with integration across topics • High quality materials, well-aligned with course goals
Teaching practices <ul style="list-style-type: none"> • How is in-class and out-of-class time used? • What assignments, assessments and learning activities are implemented to help students learn? 	<ul style="list-style-type: none"> • Teaching practices are not sufficiently planned or organized, or are poorly implemented • Practices are not well executed; little development in methods despite evidence of need • Students lack opportunities to practice the skills embedded in course goals • Student engagement is variable 	<ul style="list-style-type: none"> • Teaching practices are well planned and organized • Standard course practices carried out; follows conventions within discipline and institution • Students have some opportunities to practice skills embedded in course goals • Students consistently engaged 	<ul style="list-style-type: none"> • Activities are well planned, integrated, and reflect commitment to providing meaningful assignments and assessments • Uses effective, high-impact or innovative methods to improve understanding • In- and out-of-class activities provide opportunities for practice and feedback on important skills and concepts • Students show high levels of engagement
Achievement of learning outcomes <ul style="list-style-type: none"> • What impact do these courses have on learners? • What evidence shows the level of student understanding? 	<ul style="list-style-type: none"> • Insufficient attention to student learning—quality of student learning is not described or analyzed with clear standards • Evidence of poor student learning; low level of skill/ understanding is required or achieved without clear attempts to improve 	<ul style="list-style-type: none"> • Clear standards for evaluating the quality of student understanding • Typical student achievement for courses at these levels 	<ul style="list-style-type: none"> • Standards for evaluating student understanding are connected to program or curriculum expectations, or use authentic assessments • Efforts to support learning in all students • Quality of learning supports success in other contexts (e.g., subsequent courses or non-classroom venues), or is increasing over successive offerings
Classroom climate and student perceptions <ul style="list-style-type: none"> • What are the students' views of their learning experience? • How has student feedback informed the faculty member's teaching? 	<ul style="list-style-type: none"> • Classroom climate does not promote civility or discourages student motivation and engagement • Consistently negative student reports of teacher accessibility, interaction skills • Poor sense of learning among students • Little attempt to address concerns voiced by students 	<ul style="list-style-type: none"> • Classroom climate promotes civility • No consistently negative student ratings of teacher accessibility, interaction skills • Most students indicate progress with their learning • Instructor articulates some lessons learned through student feedback 	<ul style="list-style-type: none"> • Evidence that classroom climate is respectful, cooperative, and encourages motivation and engagement • Student feedback on teacher accessibility, interaction skills is generally positive • Students perceive that they are learning important skills or knowledge • Instructor is responsive to student feedback in short- and long-term
Reflection and iterative growth <ul style="list-style-type: none"> • How has the faculty member's teaching changed over time? • How has this been informed by evidence of student learning? 	<ul style="list-style-type: none"> • No indication of having reflected upon or learned from prior teaching or feedback 	<ul style="list-style-type: none"> • Continued competent teaching, possibly with minor reflection based on input from peers and/or students • Articulates some lessons learned from prior teaching and feedback 	<ul style="list-style-type: none"> • Regularly makes adjustments to teaching based on reflections on student learning, within or across semesters • Examines student performance following adjustments • Reports improved student achievement of learning goals based on past course modifications
Mentoring and advising <ul style="list-style-type: none"> • How effectively has the faculty member worked individually with UG or graduate students? 	<ul style="list-style-type: none"> • No indication of effective mentoring or advising students (but expected in department) 	<ul style="list-style-type: none"> • Some evidence of effective advising and mentoring (<i>define as appropriate for discipline</i>) 	<ul style="list-style-type: none"> • Evidence of exceptional quality and time commitment to advising and mentoring (<i>define as appropriate for discipline</i>)
Involvement in teaching service, scholarship, or community <ul style="list-style-type: none"> • In what ways has the instructor contributed to the broader teaching community, both on and off campus? 	<ul style="list-style-type: none"> • No interaction with broader community about teaching, including involvement with teaching-related committees • No evidence of keeping up with reports on effective teaching • Practices and results of teaching are not shared with others • Actions have negative impact on teaching culture in department or institution 	<ul style="list-style-type: none"> • Some involvement in teaching-related committees, or engagement with peers on teaching (e.g., teaching-related presentations or workshops) • Participates in department-level curriculum decisions 	<ul style="list-style-type: none"> • Regular involvement in teaching-related committees, engagement with peers on teaching (e.g., teaching-related presentations or workshops) • Occasional (or more) local or external presentations or publications to share practices or results of teaching • Contributes to department or university curricular planning or assessment • Advanced—Scholarly publications or grant applications related to teaching

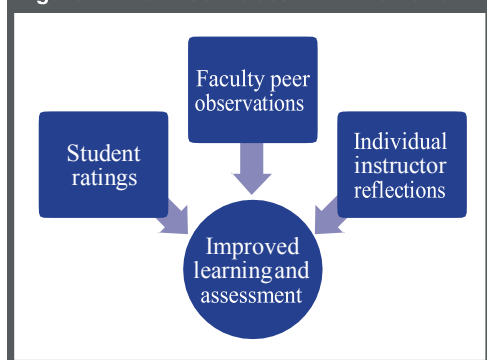
TEACHING QUALITY FRAMEWORK (TQF)

The University of Colorado Boulder requires that “[d]ossiers for comprehensive review, tenure, or promotion must include multiple measures of teaching.” (Guidelines 2007) However, at present we do not have a well-defined framework to guide individuals or departments in the selection and interpretation of such measures, which makes it difficult to assess teaching quality and support systemic faculty growth in teaching. In this project, we outline a framework for supporting and assessing teaching quality for all instructors across all departments on campus that is grounded in the scholarship of higher education. Such a framework will advance individual educational efforts as well as support the alignment of campus

The Framework

The goal of the framework is to **support improved teaching** by providing faculty members with feedback that they can use to improve as educators and to **provide better mechanisms for assessing** teaching quality for tenure, promotion, and merit.

Figure 1: The three “voices” in the framework.



This framework defines teaching as a scholarly activity (like research) and assesses core components of such scholarship. One example we draw from is: Glassick et al, *Scholarship Assessed: Evaluation of the Professoriate*, '97 (more on Pg 2).

1. clear goals,
2. adequate preparation,
3. appropriate methods,
4. significant results,
5. effective presentation, and
6. reflective critique.

This assessment of framework criteria is made through the use of the three standard “voices” (data sources): the faculty member, the students, and peers. These framework categories are held constant across all departments; however, the **definition and interpretation of these components of the framework** (making them specific) **and their relative weights would be defined at the unit level**. Thus departments specify in a clear way what is meant by “multiple measures” locally, but using common categories across campus. This approach provides the university with a common framework while preserving disciplinary identity and specificity.

The Process

The implementation of the TQF that is **not** a top-down mandate, but instead focuses on bringing together key faculty leaders and departments and providing them with a structure to help them co-create, test, and evaluate the framework. This is an opt-in model, with pilot departments choosing to engage and become leaders in this process. Thus, this strategy empowers the community to voluntarily engage in the exploration of new ways of assessing teaching and to adopt the framework because they see its value.

Departmental TQF Teams:

- 12+ Depts in A&S, Engineering, Business involved
- 3-4 leads in each department
- Tasked with contextualizing the elements of the framework to the discipline and deciding what resources and process are required for implementation in their department.
- Following a Dept. Action Team model.

Campus / Cross-Unit TQF Dialog:

- Wide participation (departmental representatives, deans, VC-level & other key stakeholders)
- Defining the TQF & including changes proposed by the departmental teams.
- Communicate with campus T&P committees, non-pilot departments, etc.

A **facilitator** will support multiple departmental TQF teams and act as a **communication channel** across the departmental teams. Additionally, we expect the departmental teams to generate lists of **required resources** necessary to make the implementation of the TQF feasible given limitations on faculty time.



University of Colorado
Boulder

in collaboration with UKansas, UMass & Mich.St. U

Contacts: Noah Finkelstein, Jessica Keating, Joel Corbo, & Mark Gammon (CU), Daniel Reinholz (SDSU), Daniel Bernstein (Univ. of Kansas)
Center for STEM Learning www.colorado.edu/csl

Ernest Boyer's publication, *Scholarship Reconsidered: Priorities of the Professoriate* (1990), has played a key role in broadening the perception of academic scholarship. Boyer defines four types of scholarship, including the scholarship of teaching. Subsequent work, *Scholarship Assessed: Evaluation of the Professoriate* (Glassick, Huber, & Maeroff, 1997), has made great strides to operationalize the assessment of all forms of scholarship in terms of six components. Adapted from *Scholarship Assessed*:

1. **Clear goals:** Does the instructor state the goals of the course/learning experience clearly? Are these goals realistic and achievable? Do they relate to important questions in the relevant field of study?
2. **Adequate preparation:** Does the instructor have an understanding of the scholarship of teaching and learning in the field? Has he or she practiced the necessary skills and gathered the necessary resources to allow for successful learning?
3. **Appropriate methods:** Does the instructor choose teaching methods appropriate to achieve the learning goals, and does he or she apply them effectively? Does the instructor modify these methods in response to changing circumstances in the classroom?
4. **Significant results:** Does the instructor achieve his or her goals? Does the instructor's work in the classroom add consequentially to the knowledge of teaching in his or her field or open up new areas for exploration?
5. **Effective presentation:** Does the instructor communicate with his or her students using suitable style, effective organization, appropriate forums, and clarity and integrity? Does the instructor communicate the results of his or her teaching to peers using the same set of criteria?
6. **Reflective critique:** Does the instructor critically evaluate his or her teaching, using an appropriate breadth of evidence? Does the instructor use this evaluation to improve the quality of future work?

Each of these six components is elaborated in more detail in *Scholarship Assessed* and has been further operationalized by others (Bernstein et al., 2010). A sample rubric is below; though many exist to draw from.

Figure 2: Rubric for Assessing Teaching as Scholarly Activity (from Bernstein)

Components	Entry into teaching	Basic Skill	Professional	Advanced
Goals of the course or other learning activity	Course/activity goals are absent, unclear, or inappropriate.	Course/activity goals are well articulated and appropriate to the course and to the curriculum.	Course/activity goals identify intellectually challenging and enduring targets and/or are especially well matched to students.	Course/activity goals identify levels of performance that represent excellence and are of interest to many stakeholders.
Preparation for the course or learning activity	Teacher is not adequately knowledgeable and/or has no background in teaching.	The teaching is based on prior scholarship in its area, including current content as well as pedagogical methods and conceptual frames.	The teacher's preparation includes broad synthesis of prior work in content as well as practice in pedagogical methods and conceptual frames.	The teacher acquires and integrates knowledge and skills drawn from the literature of multiple disciplines, both in content and pedagogy.
Methods used to conduct the teaching	No apparent rationale for teaching methods is used; there is no instructional design.	The work follows the conventions of teaching practices within its domain of discipline and institution.	The teaching takes full advantage of effective methods discussed within its discipline.	The work generates new practices that will enable others to improve or enhance their teaching.
Evidence gathered to demonstrate the impact of the teacher's work	There is no measure of student learning, or assessment methods do not match espoused goals.	There is evidence linking students' performance to espoused goals.	Student performances indicate that deep and/or broad learning is taking place.	The learning demonstrated is exemplary in either depth of learning and/or in breadth of students' success.
Communication of teaching results to others	The practice and results of teaching are kept private.	The teacher's work and students' performances are publicly accessible for others to use, to build on, and to review critically.	The teacher's reflective work has been read and adjustments in practice have arisen through the public discourse.	The teacher's work has had an impact on the practices and inquiry of many others and has contributed to related conceptual frameworks.
Reflection on the teaching and its impact on student learning	The teacher provides no indication of having reflected on or learned from prior teaching.	The teacher articulates lessons learned from reflecting on prior teaching.	The teacher has examined the impact on students' performance within a conceptual framework and adjusted practices based on reflection.	Enhanced achievement of learning goals results from reflection on evidence within a conceptual framework, or the teacher revises the conceptual framework based on student learning outcomes.



Appendix G

NC State Peer Observation Protocol and Review of Course Materials Checklist

Table 2
Course Material Checklist

Course: _____ **Instructor:** _____ **Date:** _____

Circle your responses to each of the questions and then add comments below the table.

	Exceeds expectations in all respects	Meets expectations in all respects	Meets expectations in most respects	Meets expectations in some respects	Meets expectations in few or no respects
1. Course content includes the appropriate topics	5	4	3	2	1
2. Course content reflects the current state of the field	5	4	3	2	1
3. Course learning objectives are clear and appropriate	5	4	3	2	1
4. Course policies and rules are clear and appropriate	5	4	3	2	1
5. Lecture notes are well organized and clearly written	5	4	3	2	1
6. Supplementary handouts and web pages are well organized and clearly written	5	4	3	2	1
7. Assignments are consistent with objectives and appropriately challenging	5	4	3	2	1
8. Tests are consistent with learning objectives and appropriately challenging	5	4	3	2	1
9. Tests are clearly written and reasonable in length	5	4	3	2	1
10. Student products demonstrate satisfaction of learning objectives	5	4	3	2	1

What are the strengths of the course materials? (Continue on back if necessary)

What could have been improved? (Continue on back if necessary)

Rater(s) _____

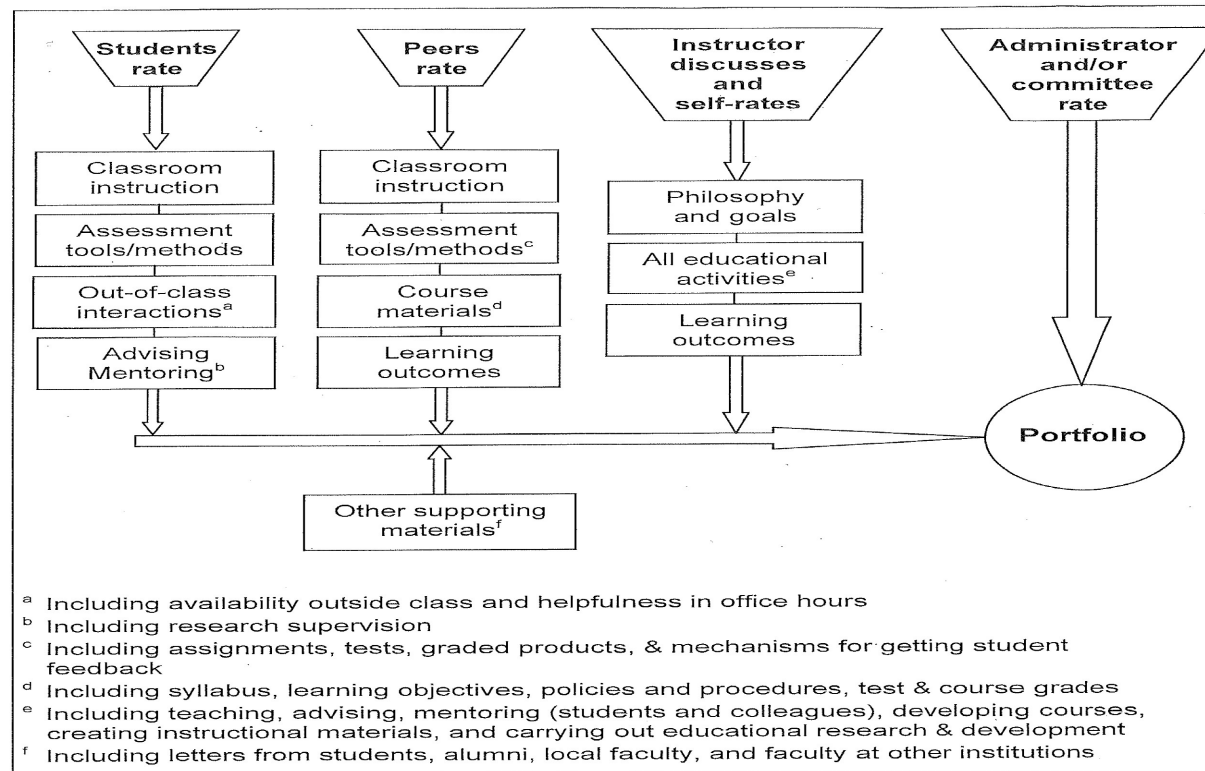


Figure 1. Comprehensive Evaluation of Teaching Performance

Appendix H
University of Oregon Peer Review Framework

Draft April 29 2018

Prepared by Senate Task Force for Teaching Evaluation

Proposed UO Peer Review of Teaching Framework

Goal:

The task force proposes that excellent teaching is inclusive, engaged and research-led, and that evidence of excellence be collected from student experience surveys, faculty course reflections and peer review reports. The goal of peer review is to provide evidence and recommendations to be used for both a) continuous course improvement and b) evaluation of teaching excellence.

Philosophy:

Peer review is a key component to the evaluation of teaching. However, some peer reviews do not provide useful feedback for a) course improvement or b) evaluation of teaching excellence, especially when there is no clear definition of teaching excellence and no clear guidelines for the reviewer. Developments in the scholarship of teaching of learning now provide the tools to produce meaningful classroom observations based on evidence-based practices. Peer review of teaching should become one of the key components in the evaluation of teaching excellence, mirroring its use in all other aspects of academic work.

Current problems to be solved:

- No definition of teaching excellence is provided to reviewer.
- Most peer observations are un-structured and vary in effectiveness for a) course improvement and b) evaluation of teaching excellence.
- Many peer observations are rushed to fulfill tenure and promotion dossiers.
- Peer observers are frequently under-trained in evidence-based, effective teaching practices.
- Faculty under review do not have the opportunity to communicate their expectations, learning goals, and pedagogical decision-making to the reviewer.

Proposed Framework:

All units **create a policy** that outlines the specific requirements for the peer review of teaching system that includes the following:

- 1) Coordinator to track departmental peer reviews
- 2) Formal and evidence-based observation tools
- 3) Faculty self-assessment tool
- 4) Structured reviewer-instructor follow-up meeting
- 5) Template for peer review report
- 6) Training for peer reviewers

The details for each requirement of the peer review system are listed below.

Draft April 29 2018

Prepared by Senate Task Force for Teaching Evaluation

1. Coordinator to track department peer reviews:

Each unit/department will develop a policy, which outlines each step in the process (aligned with this framework), as well as a **Peer Review Coordinator** who oversees the scheduling of all peer reviews for the year. The coordinator can be a faculty member, unit head, or unit manager. The peer review schedule for the year will be shared with all faculty by week 1 of Fall term so that adjustments can be made if needed/requested. Peer reviews will be scheduled based the frequency they are required for each faculty classification and rank as described below:

Pro Tem Faculty – once per year

Career Instructional Faculty – once per contract period

Assistant Professor – once per year

Associate Professor – once every other year

Full Professor – once every three years

2. Formal and evidence-based observation tool:

To ensure course observations are consistent and based on the scholarship of teaching and learning, each unit will select one course observation tool to be used for all peer reviews. If the observation tool is created by the unit/department, it requires references to the scholarship of teaching and learning in the discipline and should align with the department's vision and learning objectives. Otherwise, please select a published tool such as one of the following.

[TEP Peer Teaching Observation Guide \(customizable\)](#)

[COPUS: Classroom observation protocol for undergraduate STEM](#)

3. Faculty self-assessment tool:

The unit/department shall identify a self-assessment tool that is included in the peer review process. Self-assessment tools provide the faculty member the opportunity to reflect on their teaching practices and observe changes over time. The tool selected should provide opportunities for specific recommendations for continued improvement. If the self-assessment tool is created by the unit/department, it requires references to the scholarship of teaching and learning in the discipline and should align with the department's vision and learning objectives. Otherwise, please select a published tool such as one of the following:

[TEP Faculty Self-Assessment Guide](#)

Teaching Practices Inventory developed for [STEM](#) and [Social Sciences](#)

4. Structured reviewer-instructor follow-up meeting:

After the classroom observation and faculty self-assessment has been completed, the reviewer and faculty must meet to find out more about the faculty's inclusive, engaged and research-led teaching practices. A consistent list of questions must be outlined by the unit/department, which will form the basis of the discussion and be included in the report. Unit-developed questions require references to the scholarship of teaching and learning in the discipline and should align with the department's vision and learning objectives, or the following can be used:

- a) What specific methods do you use to ensure this course is inclusive of diverse students and scholar identities?
- b) How do you engage your students during class time? How do you engage your students outside of class time?
- c) How do you access the scholarship of teaching and learning (books, conferences, workshops, journal articles, peer observation etc.)?
- d) Name some of the research-led pedagogies that you are incorporating into the course (research-led pedagogy can include infusing your research into the course, engaging students in research, and using teaching practices described in the scholarship of teaching and learning to be most effective).

5. Template for peer review report:

By creating a template for the output of a Peer Review, the unit/department, school/college and university personnel committees can expect consistent, robust reports that provide information that is valuable for both a) continual course improvement and b) evaluation of teaching excellence. The report should include the following sub-headings:

a. Overview: Include the course name/number, time and date, and the topics under discussion that day. Include the context of the course, size of the class, type and level of students (majors/non-majors, freshmen/seniors, elective/required course).

b. Information collected: Description of the information collected from:

- i) classroom observation tool
- ii) self-assessment tool
- iii) answers to questions posed during reviewer-instructor follow-up meeting.

c. Recommendations: Based on the information collected, provide recommendations to the individual being evaluated that will continue to support student success through the use of inclusive, engaged and research-led teaching

in the context of the specific course under review. The recommendations will provide insight regarding the progress toward teaching excellence.

d. References: Provide a list of references that form the basis for the classroom observation tool, the self-assessment tool and the questions for the follow-up meeting (which will be the same for all reports from one unit/department).

6. Trained Peer Reviewers:

Each unit must identify and train a group of faculty to serve as peer reviewers. Participation should count as important unit/department level service, and typically requires 4-6 hours of service per faculty review. The unit/department could either train all faculty, or only the subset of faculty who will perform all peer reviews for the year. Faculty who will serve as reviewers will be identified at the start of each academic year. Faculty outside of the unit/department (e.g.: Teaching Academy members) may provide external peer review, provided they are trained in the unit's/department's protocol.

Selected Bibliography

- Association of American Universities. "Aligning Practice to Policies: Changing the Culture to Recognize and Reward Teaching at Research Universities." Research Corporation for Science Advancement, 2017.
- Association of American Universities. "Student Identification of Learning Outcomes and Improved Student Evaluations." *Searching for Better Approaches: Effective Evaluation of Teaching and Learning in STEM*. Research Corporation for Science Advancement, 2015.
- Bain, Ken. *What the Best College Teachers Do*. Cambridge, MA: Harvard UP, 2004.
- Basow, Susan, and Silberg, Nancy. "Student Evaluations of College Professors: Are Female and Male Professors Rated Differently?" *Journal of Educational Psychology* 79 (3): 308-14. 2013.
- Benton, Stephen and Cashin, Stephen. "Student Ratings of Teaching: A Summary of Research and Literature." IDEA Paper #50. The IDEA Center. 2011.
- Benton, Stephe, and Ryalls, Kenneth. "Challenging Misconceptions About Student Ratings of Instruction." IDEA Paper #58. The IDEA Center. 2016.
- Berk, R., & Theall, Mike. *Thirteen Strategies to Measure College Teaching : A Consumer's Guide to Rating Scale Construction, Assessment, and Decision Making for Faculty, Administrators, and Clinicians* (First ed.). Sterling, Virginia: Stylus. 2006.
- Boring, et. al. "Student Evaluations of Teaching (Mostly) Do Not Measure Teaching Effectiveness" Science Open Research, 2016.
- Boston University, Center for Teaching & Learning: Learning Analysis Poll: www.bu.edu/ctl/teaching-resources/evaluation-of-instruction/learning-analysis-poll
- Braga, Paccagnelli, and Pellizzari. "Evaluating Students' Evaluation of Professors." Institute for the Study of Labor, Bonn, Germany. 2011
- Brown University. "Report on the Student Course Evaluation Instrument." May 2018
- Carroll, Scott and James West. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors." *Journal of Political Economy*, 2010.
- Dangel, H., & Lindsay, P. "What Are Our Students (Really) Telling Us?" *Journal of Faculty Development*, 28 (2): 27-34. 2014.
- Diamond, M. R. "The usefulness of structured mid-term feedback as a catalyst for change in higher education classes." *Active Learning in Higher Education*, 5(3), 217-231. 2004.
- Ewing, Andrew, "Estimating the Impact of Relative Expected Grade on Student Evaluations of Teachers." *Economics of Education Review*, 2008.
- Falkoff, Michelle. "Why We Must Stop Relying on Student Ratings of Teaching." *Chronicle of Higher Education*. 25 April 2018.

- Hanover Research Council. "Best Practices in Student Course Evaluation." 2009.
- . "Best Practices in Using Aggregate Course Evaluation Data." 2014.
- Kaplan, M. *The Teaching Portfolio*. Occasional Paper No. 11. Ann Arbor, MI: Center for Research on Learning and Teaching, University of Michigan. 1998.
- Linse, Angela. "Interpreting and Using Student Ratings Data: Guidance for Faculty Serving on Evaluation Committees." *Studies in Educational Evaluation* 54: 94-106 (2017).
- McGowan, W. R., & Osguthorpe, R. T. Student and Faculty Perceptions of Effects of Midcourse Evaluation. *To Improve the Academy*, 29 (1), 160–172 (2011).
- McPherson, Jewel, and Kim. "What Determines Student Evaluation Scores? A Random Effects Analysis of Undergraduate Economics Classes." *Eastern Economic Journal* (2009).
- Meizlish, D., & Kaplan, M. "Valuing and evaluating teaching in academic hiring: A multi-disciplinary, cross-institutional study." *Journal of Higher Education* 79(5), 2008.
- Mengel, Friederike, Sauermann, and Zölitz. "Gender Bias in Teaching Evaluations." *Journal of the European Economic Association*, 2017.
- Mitchell, Kristine, and Martin, Jonathan. "Gender Bias in Student Evaluations." *Political Science and Politics* 51 (5), 648-52, 2018.
- Ragupathi, K. Resource Guide for Faculty: Gathering formative feedback through mid-semester evaluations. Singapore: National University of Singapore (n.d.).
- Seymour et. al., "Creating a Better Mousetrap: Online Student Assessment of their Learning Gains" Presented at the National Meetings of the American Chemical Society Symposium, 2000.
- Stanford University. Course Evaluation Committee Report. 2013
- Stark, Philip and Joseph Freishtat. "An Evaluation of Teaching Evaluations." *Science Open*. 2014.
- Supiano, Becky. "What Professors Can Learn About Teaching from Their Students." *Chronicle of Higher Education*. 19 November 2017.
- Vlieger, Jacob, and Stange. "Measuring Up: Assessing Instructor Effectiveness in Higher Education." *Education Next*, 2017.
- Wagner et. al. "Gender, Ethnicity, and Teaching Evaluations 2016: Evidence from Mixed Teaching Teams." *Economics of Education Review*, 2016.
- Weinberg, Bruce, Fleisher, Belton, and Hashimoto, Manasori. "Evaluating Methods for Evaluating Instruction: The Case of Higher Education." National Bureau of Economic Research, 2007.
- Wieman, Carl. "A Better Way to Evaluate Undergraduate Teaching." *Change Magazine*, 2015

Wieman, Carl and Gilbert, “The Teaching Practices Inventory: A New Tool for Characterizing College and University Teaching in Math and Science.” *CBE Life Sciences Education*. Fall 2014.

Wohlfarth et. al., “Student Perceptions of Learner-Centered Teaching.” *Insight: A Journal of Scholarly Teaching*, 2008.