Recommendations for the Future
Evaluation of Teaching Effectiveness at Penn State
July 2020
Committee on Assessing Teaching Effectiveness

Members:
Cynthia Brewer, Professor and Head, Department of Geography
Melina Czymoniewicz-Klippel, Assistant Teaching Professor, Biobehavioral Health, College of Health and Human Development
Julie Gallagher, Associate Professor of History and American Studies, Penn State Brandywine
Keith Gilyard, Edwin Erle Sparks Professor of English and African American Studies and Director of Graduate Studies, Department of African American Studies, College of the Liberal Arts
Tim Kurzweg, Professor, Electrical and Computer Engineering, and Director, School of Engineering, Penn State Behrend
Angela Linse (chair), Associate Dean and Executive Director, Schreyer Institute for Teaching Excellence

Introduction

This is the second report of the Committee on Assessing Teaching Effectiveness submitted to Kathy Bieschke, Vice Provost for Faculty Affairs. One part of the committee's charge was to investigate options for improving future evaluation of teaching for tenure, promotion, annual review, and reappointment. This report addresses the unacceptable over-reliance on student feedback, specifically the Student Ratings of Teaching Effectiveness (SRTE) numerical ratings and 'Open Ended Item' responses, which serve to amplify systemic inequities and hierarchies within our teaching community.

We recognize that effective teaching rests fundamentally on the intentional design and implementation of teaching practices that support student achievement of course-level learning outcomes. [1]

Below we recommend numerous options for faculty to document and for departments and/or administrators to evaluate teaching effectiveness to include in promotion dossiers and review materials. These options include a faculty member's self-reflection on their teaching, peer and/or discipline specific input, and student feedback. The recommendations are crafted to decenter the personal terms that students frequently use to describe and evaluate the faculty member, and to emphasize the teaching and learning.

Rationale for this Report

We offer these recommendations at this critical juncture for several reasons, including concerns about bias, the impacts of the COVID -19 pandemic, and a long overdue review of teaching evaluation policy. The publication of the "More Rivers to Cross" report documents the clear, disproportionate negative impacts of SRTEs on Black faculty members. The continued use of student ratings is questionable due to systemic biases that privilege white, male faculty over female faculty and faculty of color as well as the over-reliance on and misuse of student ratings data in personnel decisions. In addition, the global pandemic that hit the United States in March 2020 necessitated an immediate conversion of all face-to-

---

[1] Adapted from Developing and Evaluating Teaching Effectiveness at Colorado State University, The Institute for Learning and Teaching (TILT), Colorado State University, 2019, v. 1.1.

face courses to remote delivery. In World Campus courses, faculty and students experienced negative impacts of the pandemic as people grappled with family illness and economic fallout. The changes in course delivery necessitated by the pandemic will continue for the foreseeable future, as will the disruption of typical evaluations of teaching by peers and students. Finally, the committee determined that a review of the evidence considered for the evaluation of teaching would be beneficial given that the last review of this was conducted by University Faculty Senate in 2006[2].

Below, the committee identifies ways that faculty, peers, and students can contribute evaluative feedback on the teaching and learning experience, particularly during a period of disruption when traditional processes are not possible. The disruption offers an opportunity for the university to reconsider its (over)reliance on narrow range of evidence in the evaluation of teaching.

Recommendation 1: Faculty Self-Assessment of Teaching
Emphasis on peer review and student ratings, overshadows the voice of the instructor in describing their teaching. The options below explicitly add that third and important voice to the evaluation process, aside from the Narrative Statement in standard dossiers.

a. **Teaching Philosophy/Statement.** A teaching statement or teaching philosophy is a description of what a faculty member does to enable student learning and includes examples of how the instructor enacts that statement through course design and instruction. An effective teaching philosophy discusses students, not solely content or the faculty member's beliefs about teaching and learning. Teaching philosophies vary widely and across disciplines (examples from the Schreyer Institute for Teaching Excellence).

b. **Course Objectives**. Faculty members may reflect on how students are provided with opportunities to demonstrate achievement of course objectives and/or program outcomes.
   i. How students' work on an assignment or activity, or performance on exam or quiz items demonstrates that the assignment, activity, exam or quiz item is designed to produce evidence of learning-objective achievement.
   ii. Analysis of grading rubrics to reflect alignment between grading criteria and learning objectives.
   iii. Assessment scores (e.g. assignment grades) relative to student learning objectives.
   iv. Examples of student work (by grade level or quality rank), including grading rubrics used to score the work.
   v. In programs that are accredited through an external body, specific course outcomes for a course may be determined by a program assessment committee. These course outcomes are measured through developed metrics, such as specialized grading rubrics or assessment scores, developed by the program assessment committee to provide evidence of meeting overall program outcomes. For accredited programs that have course-level accreditation outcomes, the instructor might present measured data that support that the course objectives were met.

---

[2] Interim Report of the Special Subcommittee to Assess the Nature of Evidence Used for Promotion and Tenure Decisions, Informational Report, Senate Committee on Faculty Affairs, University Faculty Senate, Penn State, March 14, 2006, Appendix J.

c. **Intellectual Work of Teaching.** Faculty may reflect on the expertise involved in teaching their courses, citing specific examples. Example topics that might guide this reflection:
    i.   Course planning that includes content knowledge, selection of sources, and anticipation of students' prior learning or misconceptions, and strategies for overcoming disciplinary bottlenecks.
    ii. Creating connections to research in the field and/or professional practice.
    iii. Course design that creates learning pathways or scaffolding through deliberate links between assignments, readings, lectures, discussions, and/or other course elements.
    iv. Course changes in response to pedagogic innovation, student learning needs, or remote learning modes.
    v.   Examples, with quantitative or qualitative evidence, of how the instructor has demonstrated continual course improvement using varied feedback mechanisms such as ungraded [Classroom Assessment Techniques](), a [mid-semester class interview]() or a [student survey]().

Recommendation 2: Peer Assessment of Teaching
Expert teachers are a good choice for evaluating teaching abilities of other instructors and there are a variety of ways the voices of peer evaluators may be included in promotion dossiers, reviews, and reappointment materials.

a. **Peer Observation:** Peer faculty members who sit in on an in-person class, synchronous online session, or an asynchronous module of an asynchronous online course are encouraged to examine the content of the course syllabus and schedule and assess whether the reviewed session or module aligns with the course objectives and how it fits into the course design. This may be accomplished by adding the peer to a Canvas course for a short period of time to review the material. The observer may also review course content produced during or after the observed session or module, such as online posts, discussions, responses, or homework. This is a standard practice in peer assessments of online teaching, but is also important for in-person, hybrid, remote, and other mixed-mode offerings.

    In-class peer observation best practices include planned visits and being invited by the instructor, in consultation with the head of the academic unit. Another characteristic of effective peer observation is a focus on what the instructor wants to learn or improve about their teaching, which necessitates a conversation in advance of the observation. Finally, faculty within an academic unit should consider coming to a consensus about what constitutes effective teaching in order to ensure that reviewed instructors are receiving a consistent message. One way to achieve consistency is to use a peer observation template or rubric.

    For early career faculty, peer observations should be conducted by a faculty member of higher rank. For faculty of the middle and highest ranks, reviewers may be of equivalent or higher rank. Academic administrators may also conduct observations of teaching performance and materials.

    Peer and administrator observations may be reported in a letter format made available to the instructor at the time of the review, and included in review or promotion materials; i.e., these should be considered confidential evaluations.

b. **Peer Review of Course Materials** (peer evaluates a packet or portfolio for a single course): The faculty member under review creates a compact portfolio of course materials to be reviewed by a peer. The course material should reflect the learning objectives and select materials that demonstrate how the instructor enables student success. Documentation might include the

syllabus, readings, a sample of key topics, assignments, lab experiences, discussion prompts, and assessments such as quizzes or exams. This compilation may be introduced with narrative about the course and how the materials are related as well as a table of contents. Example student products may also be included, but student privacy must be maintained and all identifiable information redacted.

Selection of the peer evaluator is flexible, not hierarchical, and should be a peer from another Penn State campus or a faculty member who has similar expertise or teaches a similar course. Reviewer selections should be made jointly by the instructor and administrator or supervisor and the evaluation may be requested by either. Reviewers should be given at least a month to the review the course material and return a review letter assessing the relevance of the course and whether the design aligns with the course objectives. Reviewers should be requested to maintain the confidentiality of instructor's intellectual property. It is the evaluator's letter (not the portfolio) that becomes part of the promotion dossier, annual review, or reappointment package.

As above, an administrator may also conduct a review of the course material. This may replace or add to an in-class observation.

c. **Peer Review of 'Teaching Portfolio'** (referencing materials from a range of courses): Faculty members present a portfolio that reflects the breadth of the instructor's teaching. This option invites evaluation by a peer or external writer. Teaching portfolios typically accompany a teaching philosophy statement and course materials provide evidence that supports the statement. The portfolio provides a larger context about the instructor's teaching and includes references to how their courses relate to program goals and requirements, to each other, and to other instructors' offerings.

    Reviewers chosen to evaluate a teaching portfolio should be experts in the instructor's field(s) who have an understanding of the curriculum in a discipline. The reviewer might be a Penn State faculty member at another campus or a peer at another institution or organization (such as a federal agency). The reviewer provides a letter of evaluation that becomes part of the external letters section of a promotion dossier. For tenure-line faculty, including at least one external letter of teaching evaluation may address a common perception that Penn State values research above teaching.

    Selection of reviewers should be made jointly by the instructor and administrator or supervisor, and the evaluation letter is requested by the administrator (e.g., program head). The consultation is important because the instructor (or supervisor) in order to protect the instructional intellectual property of the faculty member or a unique contribution to the program. External reviewers must be trusted to respect the boundaries of intellectual property, particularly with specialty and/or online programs. The confidentiality issue is different for teaching materials because course materials are not publicly available as are research, scholarship, or creative products. Reviewers should be given at least 2 months to complete the review and return a letter of evaluation of the portfolio.

d. **Course and Program Accreditation Outcomes:** Accreditation is a voluntary, nongovernmental process that includes an external review to identify quality programs (e.g., Accreditation Board for Engineering and Technology (ABET), Association to Advance Collegiate Schools of Business International (AACSB)). The accreditation process ensures that students are learning material most relevant to their field of study, preparing them to be effective leaders upon graduation.

Included in the accreditation process is a comprehensive review of a school's mission, faculty qualifications, and the curricula, including courses and their support of the overall program. In some, specific course outcomes are identified, and assessment metrics are created to measure the success and effectiveness of the teaching and student learning.

In academic programs accredited from an external organization, a faculty member's ability to meet the program objectives can be evaluated using specific accreditation metrics. Typically, metrics can be measured through specialized grading rubrics or assessment scores. This peer review could be conducted by an experienced external reviewer, a program or department head, or a faculty member familiar with the accreditation objectives and standards.

Recommendation 3: Assessment by Students
As part of the on-going effort to revise the criteria for teaching effectiveness, we propose the following options for evaluation of instructors by students.

a. **Advising/Mentoring**: Offer students the opportunity to provide feedback about their experience of being mentored and/or advised by the instructor.

b. **Surveys of previous students or alumni:** Conduct surveys of students at an interval from two- to five-year after taking a course. The survey should invite reflection on how the course and instructor influenced or prepared them for other courses or their current context.

c. **Written Student Feedback** (that is gathered simultaneously with the SRTEs)
The committee recommends implementation of both of these suggestions.
   i. Eliminate use of student all open-ended feedback in the evaluation of teaching (summative assessment), including responses to the university open ended questions and any open-ended Additional Questions written by the academic unit.

   Not only does the written feedback double-count each student's input, current methods of summarizing the student written comments are inconsistent across and within campuses and colleges. More problematic is the common practice of including negative comments even when those comments to not reflect the collective views of students. When a few negative comments are included, it gives greater weight to negative comments.

   The written comments should continue to be collected because they can provide insight on common student concerns and aid in the interpretation of SRTE items.
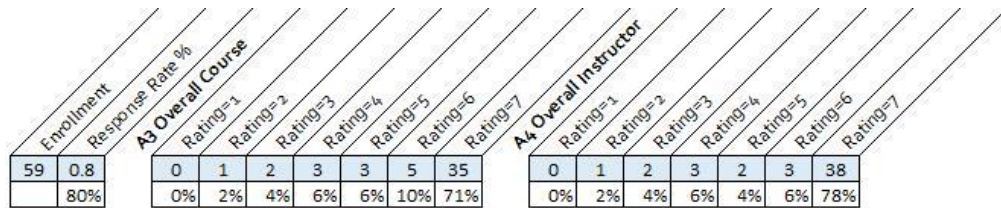
   ii. Replace the current open-ended questions more explicit questions or guidance about what constitutes appropriate subjects for student commentary. This is intended to set parameters and reduce personal comments about the instructor. For example:

   - Students could be asked to reflect on the specific course elements that helped them learn (e.g., syllabus, learning objectives, readings, power points, films, small group activities, labs).
   - Shift the emphasis to comments about class more broadly by asking: What did **we** do in this course that helped your learning? What could **we** change to improve your learning?
   - Shift the focus on the student. What did **you** do in course that helped you learn? What could **you** change to improve your learning?

d. **SRTE Recommendations:**
   i. Replace the use of SRTE quantitative data with a faculty member's written summary and reflective response to them (e.g., [Sample Student Ratings Annotation](#)). This option encourages faculty to focus on course and instructional improvements and eliminates the burden of interpretation for administrators or peers.

   If quantitative SRTE data continue to be used, all of the remaining changes should be implemented:

   ii. Eliminate the mean rating. Provide only the distribution across scores (1–7) and include the full distribution in dossiers and Activity Insite reports. For example:

Enrollment	Response Rate %	A3 Overall Course	Rating=1	Rating=2	Rating=3	Rating=4	Rating=5	Rating=6	Rating=7	A4 Overall Instructor	Rating=1	Rating=2	Rating=3	Rating=4	Rating=5	Rating=6	Rating=7
59	0.8		0	1	2	3	3	5	35		0	1	2	3	2	3	38
	80%		0%	2%	4%	6%	6%	10%	71%		0%	2%	4%	6%	4%	6%	78%

   The above is similar the distributions already included in the individual SRTE summary reports, except without the graphic element, which might be problematic to import into Activity Insight:

		Lowest Rating		Average Rating			Highest Rating		N	Mean	N/A	
		1	2	3	4	5	6	7				
A3	Rate the overall quality of this course.	0/ 0%	0/ 0%	2/13%	2/13%	4/27%	3/20%	4/27%	15	5.33	0	1234567
A4	Rate the overall quality of the instructor.	0/ 0%	1/ 7%	1/ 7%	2/13%	2/13%	3/20%	6/40%	15	5.53	0	1234567

   iii. Immediately eliminate all known problematic questions from the existing [177 SRTE items](#), including items likely to prompt implicit biases (e.g. questions about authority, mastery, or presence), compound items (using the terms "and" or "or"), items that are not actionable or under the control of the instructor, poorly worded items, and out-of-date items.

   iv. Consider restricting the availability of existing SRTE forms (associated with previously administered SRTEs) to prompt academic units to reconsider their choice of items, some of which have not changed in decades. Item choices should reflect current instructional values and norms and all items should be reviewed to ensure they do not amplify stereotypes or bias or unnecessarily trigger negative reactions.

Recommendation 4: Professional Development for Teaching (instructional improvement)
This committee strongly recommends addition of a new section to all promotion dossier and evaluation packets. Inclusion of this section appropriately emphases instructional improvement and recognizes ongoing efforts to learn about effective teaching practices and behaviors and accumulate knowledge about how students learn.

Examples of professional development activities that might be included:
- Participation in faculty teaching communities or communities of practice.
- Consulting with college or campus instructional designers, Schreyer Institute for Teaching Excellence faculty, Teaching and Learning with Technology (TLT), librarians, or multimedia staff.

- Independent work through readings, webinars, virtual teaching conferences, etc.
- Participation in teaching journal or book groups
- Curricular revisions and instructional innovations, including resources for students.
- Development of instructional resources or materials for other instructors or graduate student TAs or instructors.
- Teaching and learning presentations for other instructors or at conferences (virtual and in-person).
- Collaborations with other instructors on teaching and learning projects.
- Service on teaching and learning, curriculum, and program assessment committees.
- Teaching and Learning Scholarship (course- or curriculum-based research) that results in scholarly output such as publications and conference presentations.
- Formal and informal out-of-class learning activities (e.g., undergraduate research, students-as-partners scholarship).