Work in Progress: Teaching Evaluation Demonstration Project

Dr. Vicki V. May, P.E., Dartmouth College

Vicki V. May, Ph.D., P.E. is an Instructional Professor of Engineering and the Engineering Education Program Area Lead at the Thayer School of Engineering at Dartmouth. Her research focuses on engineering education and K-12 outreach. She teaches courses in solid mechanics, structural analysis, and design. Prior to relocating to the east coast, Professor May was an Associate Professor of Architectural Engineering at the California Polytechnic State University in San Luis Obispo.

Prof. Petra Bonfert-Taylor, Dartmouth College

Petra Bonfert-Taylor is the Associate Dean for Diversity and Inclusion and a Professor and Instructional Designer at the Thayer School of Engineering at Dartmouth College. She received her Ph.D. in Mathematics (summa cum laude) from Technical University of Berlin (Germany) in 1996 and subsequently spent three years as a postdoctoral fellow at the University of Michigan before accepting a tenure-track position in the Mathematics Department at Wesleyan University. She left Wesleyan as a tenured full professor in 2015 for her current position at Dartmouth College. Petra has published extensively and lectured widely to national and international audiences. Her work has been recognized by the National Science Foundation with numerous research grants. She is equally passionate about her teaching and has recently designed and created a seven-MOOC Professional Certificate on C-programming for edX for which her team won the ”2019 edX Prize for Exceptional Contributions in Online Teaching and Learning”. Previously she designed a MOOC ”Analysis of a Complex Kind” on Coursera. A Fellow of the Association of Women in Mathematics, the recipient of the New Hampshire High Tech Council 2018 Tech Teacher of the Year Award, the Binswanger Prize for Excellence in Teaching at Wesleyan University and the Excellence in Teaching Award at the Thayer School of Engineering, Petra has a strong interest in broadening access to high-quality higher education and pedagogical innovations that aid in providing equal opportunities to students from all backgrounds. In her role as Associate Dean for Diversity and Inclusion Petra is working on initiatives that prioritize access and inclusion, improve the wellbeing of the community and create a more equitable future.

Eugene Korsunskiy

Work in Progress: Teaching Evaluation Demonstration Project

Background

Currently, evaluations of teaching at the Thayer School of Engineering at Dartmouth (Thayer) are driven by end-of-course student course evaluations. Student evaluations are used during annual reviews, as well as at reappointment, tenure, and promotion points. Attention is paid primarily to the following two Likert-scale questions from these evaluations: "I think the overall quality of the course was…" and "I think the overall effectiveness of the teaching was…". Student responses to these vague questions have been shown to be biased, and more indicative of students' perception of the instructor's charisma than any observable best teaching practices (Wieman, 2015); over-reliance on student evaluations is particularly damaging to faculty from underrepresented backgrounds (O'Meara, et al., 2022). In addition to student course evaluations, letters from students are solicited at tenure and promotion time. Finally, classroom visits are performed by a senior faculty member but again only at tenure and promotion time. On an annual basis only student evaluations (two numbers from the evaluations) are regularly used to evaluate teaching.

Inconsistency and a lack of clear guidelines, expectations, and requirements are challenges of the current system. No standard process is followed across every candidate's promotion process, and it is difficult for candidates to anticipate their own standing. Strong consensus exists among our faculty that the current process to evaluate teaching is inadequate.

In the spring of 2021, our new Dean convened a working group charged with revamping the evaluation of teaching. The group's main charge was to design a more equitable, consistent, and transparent way to evaluate teaching, starting with the integration of formative feedback during a faculty member's early career stages and tying in with our newly developed and evolving faculty mentoring program. In addition, we are designing a multi-pronged feedback system that will help instructors reflect on their teaching and receive the support they need in order to improve their teaching continuously. Our new evaluation system will consist of three main parts: student impressions, self-reflections, and peer observations.

In the summer of 2022, our school was awarded a grant through AAU to participate in their AAU STEM Department Project on Teaching Evaluation. A team from Thayer is participating in an AAU Learning Community around Teaching Evaluation (AAU, 2022).

Goal and Objectives

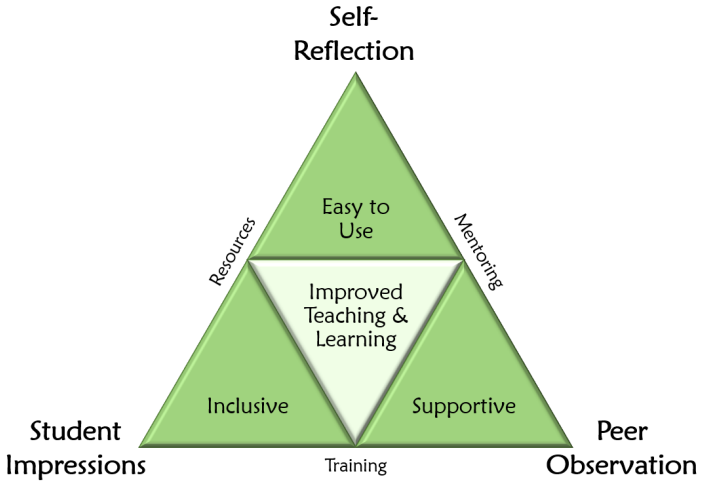
The goal of our project is to bring more consistency, clarity, and equity to the teaching evaluation process. We also hope to increase the use of evidence-based and inclusive teaching practices. Our more specific objectives and corresponding evidence of success are provided in the following table.

Objectives	**Evidence of Success**
Increase the use of evidence-based teaching practices by faculty.	Reporting and observation of the use of evidence-based teaching practices.
Involve all faculty in mentoring focused on teaching and learning.	Notes, discussions, and observations: dates, amount of time, location, topics. All faculty invited to participate.
Encourage reflection and discussion of teaching and learning among faculty.	Participation rates for completing self-reflection forms. Evaluation of self-reflections and discussions.
Increase the sense of inclusion and belonging among students.	Student evaluation responses and pre/post surveys. Increases in reported use of inclusive teaching practices.
Increase student learning and engagement.	Student responses on evaluations and observations.

Increase consistency and transparency of the teaching evaluation process.	Attitudes conveyed through Teaching Conversations, written self-reflections, and faculty meetings.

Approach

Our proposed three-pronged approach to teaching evaluation consists of instructor self-reflections, peer observations, and student impressions as outlined below and depicted at the right. Recent guidelines recommend using an evaluation approach that considers multiple perspectives (Krishnan et al., 2022).



Instructor Self-Reflection: Our Instructor Self-Reflection form includes a simplified version of the Teaching Practices Inventory developed by Wieman to encourage the use of evidence-based teaching practices in STEM courses (Wieman and Gilbert, 2014). After each course taught, an instructor is prompted to complete a self-reflection form. Some of the reflection questions are aimed at prompting the instructor to consider the extent to which their class uses evidence-based and inclusive teaching practices, such as alternative teaching modes to lecturing or cultivating a sense of belonging. The instructor also notes areas in which their course design could be improved, as well as practices that they hope to take into their next offering of the course. Instructor self-reflection forms the basis of mentoring conversations that will occur regularly, and during which additional resources are identified to help the instructor with their continuous improvement (such as workshops, connections to other instructors who may have already solved similar issues, help from IT services, etc). Finally, the next time the instructor approaches a term during which they teach the course again (at our school, it is customary to teach the same courses year after year), they will receive an automatic reminder of their previous reflections and improvement ideas.

Peer Observation: We developed a standard protocol to be used by our faculty when they visit and observe each other's classes. The protocol consists of a guided pre-observation conversation, an observation log and a few observation questions, and a set of post-observation questions. We considered adopting the Classroom Observation Protocol for Undergraduate STEM (COPUS; Smith et al., 2013) but decided it was too big of a time commitment for observers, at least initially, so developed a scaled-down version that includes an observation log instead of a minute-by-minute activities matrix. During the pre-observation conversation the observer is encouraged to learn about the class, the students, how the class fits into the arc of the term, as well as specific feedback the instructor is seeking from the observer. An observation log helps the observer focus on teaching and learning activities and student engagement, with a few follow-up questions to help the observer reflect on inclusion and other important aspects of a class session. During a guided follow-up conversation, the instructor and observer share their impressions, ideas, questions, and suggestions. We have piloted the peer observation protocol with several faculty members; feedback on the process has been very positive. In future years we envision a faculty member either selecting the peer to observe their class or having an observer assigned to them.

Student Impressions: Student feedback at Thayer and Dartmouth is sought via end-of-term course evaluations. We cannot change the current evaluation form, but we can add questions. A set of new questions has been added to the course evaluations of all Thayer faculty members. These questions focus on tangible and observable course characteristics that can be more objectively judged by students, thus eliminating some of the bias inherent in student evaluations (Boring et al., 2016, Falkoff, 2018; and Flaherty, 2019). The questions that we have been piloting include four Likert-scale questions:

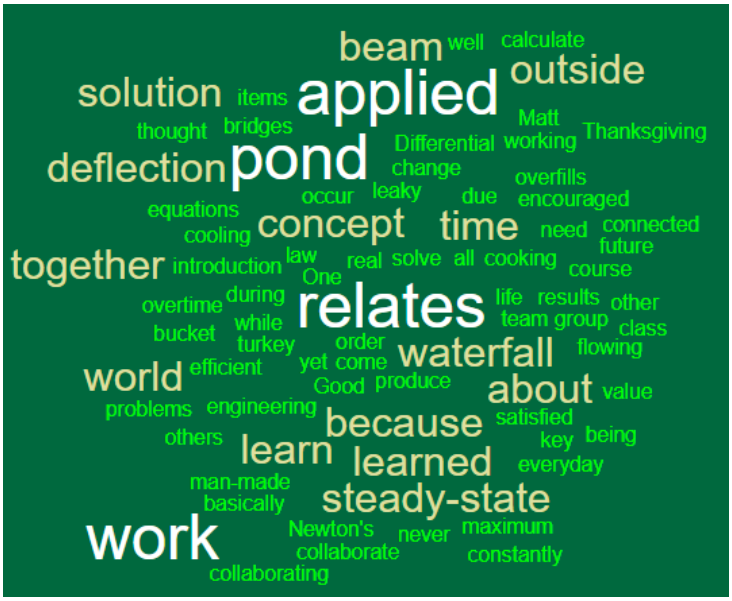
1. The instructor created and maintained an environment that was welcoming, inclusive, and respectful of diverse students and points of view.
2. The instructor provided feedback on my work in a way that helped me improve my performance.
3. The instructor provided meaningful opportunities for me to ask questions, work on problems or examples, or reflect on course material during class.
4. The instructor facilitated opportunities for me to work in groups or to collaborate with other students.

We also added one open-ended question: Describe a key concept from the class and how it relates to the world outside this course.

Pilot Results

We piloted our instruments and student evaluation questions with 16 faculty members, teaching 9 different courses in 2021-2022 and with 11 faculty members, teaching 12 different courses this year (2022-2023).

Student impressions: We are generally seeing better scores on the new, behavior focused student evaluation questions than on the two questions currently used in annual reviews (focused on overall quality and effectiveness). The average Likert rating (1=excellent and 5=poor) was 1.3 for the new questions and 1.5 for the original questions across the 16 courses evaluated in 2021-2022. A t-test shows that this is a statistically significant difference ($p=0.008$). While better teaching evaluation scores is great, the more important thing is that the instructor is receiving better feedback from the students since the new questions are focused on behaviors rather than overall perceptions. For the open-ended question (students were asked to describe a concept from the course), we've been using a word cloud to convey the responses as shown at the right (created using worditout.com). Faculty members have reported that the word clouds are a helpful way to quickly see what students are learning.



Self-reflections: All 16 of the pilot faculty participants completed the self-reflection form (sent as a google form) for their courses. They reported that completing the self-reflection form took about 15 minutes and was useful. Feedback on specific questions has been used to refine and the update the form, which is still evolving. One thing we are working on is how to adapt the form for different types of courses (lab-based and project-based, in particular).

Peer observations: Peer observations have been particularly well received by the pilot group of faculty members. Here are some quotes from faculty who have participated:

- "[I] enjoyed this more than I thought I would."
- "Appreciated the debrief: nice brainstorming that happened."
- "I actually learned stuff!"
- "Get to steal ideas."
- "It's really fun to see others teach."

- "Thank you so much for putting this all together. What a nice opportunity."
- "It's really not that much work!"

Questions

During the lightning talk, we will share our current instruments for Instructor Self-Reflection and Peer Observation. Feedback is appreciated. Questions that we have include:

- Thought on the instruments? Do they include key information without being too onerous to complete?
- What approaches have others used to improve teaching evaluation?
- How might we increase the use of evidence-based and inclusive teaching practices?
- Do campuses have different systems for annual reviews and tenure/promotion reviews?
- How is student feedback incorporated into teaching evaluations?

Future Work

Our Dean plans to integrate aspects of the new teaching evaluation process into annual faculty activity reports, which are used to document performance and determine faculty merit raises. These reports include self-reported and database collected information such as: publications, grants, course enrollment, course evaluation summaries, advisee count, committee participation, etc. New sections will be included for teaching self-reflections, teaching improvement plans and progress, and other teaching evaluation outcomes. To date, much of our work has focused on formative feedback and the annual review process. Moving forward we plan to also develop processes for incorporating the teaching evaluation process into tenure and promotion reviews. We are also working on the development of a rubric, which is based on one created by Kansas University (Follmer et al., 2020), to make it easier for the Dean to assess the teaching effectiveness of a large number of faculty members.

References

AAU (2022). "AAU to Establish Learning Communities on Effective Teaching Evaluation." https://www.aau.edu/newsroom/press-releases/aau-establish-learning-community-effective-teaching-evaluation.

Boring, A., Ottoboni, K., & Stark, P.B. (2016). "Student evaluations of teaching (mostly) do not measure teaching effectiveness." *ScienceOpen Research.* https://www.scienceopen.com/document/read?vid=818d8ec0-5908-47d8-86b4-5dc38f04b23e

Falkoff, M. (2018, April 25). "Why we must stop relying on student ratings of teaching." *The Chronicle of Higher Education.* https://www.chronicle.com/article/Whay-We-MustStop_Relying-on/243213

Flaherty, C. (2019, May 20). "Fighting gender bias in student evaluations of teaching, and tenure's effect on instruction." *Inside Higher Ed.*

Follmer Greenhoot, A., Ward, D., Bernstein, D., Patterson, M. M., & Colyott, K. (2020). *Benchmarks for Teaching Effectiveness*. (Revised 2020). https://cte.ku.edu/sites/cte.ku.edu/files/docs/KU%20Benchmarks%20Framework%202020update.pdf

Krishnan, Sandhya, Gehrtz, Jessica, Lemons, Paula P., Dolan, Erin L, Brickman, Peggy and Andrews, Tessa C. (2022). "Guides to Advance Teaching Evaluation (GATEs): A Resource for STEM Departments Planning Robust and Equitable Evaluation Practices." *CBE—Life Sciences Education*: 21:ar42, 1–15, Fall 2022 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9582816/

O'Meara, Kerryann, Templeton, Lindsey, Culpepper, Dawn, and White-Lewis, Damani, 2022. Translating Equity-minded Principles into Faculty Evaluation Reform. American Council on Education. https://www.acenet.edu/Documents/Equity-Minded-Faculty-Evaluation-Reform.pdf

Smith MK, Jones FHM, Gilbert SL, and Wieman CE., (2013). "The Classroom Observation Protocol for Undergraduate STEM (COPUS): a New Instrument to Characterize University STEM Classroom Practices." *CBE-Life Sciences Education*.

Wieman, Carl (2015). "A Better Way to Evaluate Undergraduate Teaching," *Change: The Magazine of Higher Learning*, 47:1, 6-15, DOI: 10.1080/00091383.2015.996077

Wieman, Carl and Gilbert, Sarah (2014). "The Teaching Practices Inventory: A New Tool for Characterizing College and University Teaching in Mathematics and Science." *CBE-Life Sciences Education, Vol 13(3), pp. 552–569 (2014). https://doi.org/10.1187/cbe.14-02-0023*.