

UFC/ALC Joint Task Force Members

Report on Recommendations to Establish Policy on Course Evaluations for Indiana University

April 2020

Alex Barsi Lopes (Business, IUB), Mary Fischer (Psychology, IUN), Jay Gladden (Associate Vice Chancellor for Undergraduate Education, IUPUI-Ex officio), Dennis Groth (Vice Provost for Undergraduate Education, IUB-Ex officio), Barbara Kutis (Fine Arts, IUS), Gin Morgan (Psychology, IUK), Carolyn Schult (Psychology and University Center for Excellence in Teaching, IUSB), Denise Slayback-Barry (Biology, IUPUI), Rebecca Torstrick (Senior Assistant Vice President for University Academic Affairs, Co-chair), Stephanie Whitehead (Criminal Justice and Center for Teaching and Learning, IUE; Co-chair)

Charge to Taskforce August 29, 2019

In Fall 2019, the University Faculty Council established a joint task force with the Academic Leadership Council to explore alternatives to the current student course evaluation system at IU, as part of President McRobbie's State of the University charge in 2017 to critically examine teaching initiatives at the university.

The task force was charged with exploring the development and university-wide deployment of models of student ratings of instruction that are designed to assess the learning that occurs in classes. In view of recent research that questions the value of current student evaluations of teaching, the potential for bias in the results, and a decline in narrative feedback with online evaluations, the task force was also charged to provide recommendations for improvements in the way that course evaluations are used in reviewing instructors' performance.

The task force is asked to complete its work by the end of the spring 2020 semester.

Rationale for the Work

While universities have long relied on student feedback to evaluate teaching effectiveness, faculty continue to debate the merits of these student evaluations of teaching (SETS) (Vasey and Carroll, 2016; Stark and Freishtat, 2014). Course evaluation questions often focus on student satisfaction with the course or the instructor and miss out on opportunities to assess how engaged students are on campus or in their disciplines or whether they feel they belong in college (measures which can impact student success in academic pursuits). Evidence has begun to accumulate that such feedback may not be a reliable way to judge either the efficacy of our teaching practices nor the quality of student learning in courses. One recent study (Uttl, White, and Gonzalez 2017) noted that students do not learn more from professors with higher teaching evaluation ratings. Carpenter et al. (2020) cited evidence that student *judgments* of their own learning are inflated by "learning experiences that minimize effort and increase the appearance of fluency, engagement, and enthusiasm." These findings suggest that asking students to reflect

on practices that support their learning may actually “encourage teaching practices....that do not enhance--and may even undermine--student learning...” A second study (Esarey and Valdes, 2020) was able to show using simulation that even under ideal circumstances (assuming SETS are moderately correlated with teaching quality, highly reliable, and not discriminating on any basis such as race or gender), SET scores failed to reliably identify the best teacher in a pairwise comparison and more than 25% of faculty with evaluations at or below the 20th percentile were actually above the median in instructional quality. The authors of that study argue that it is better to use multiple (imperfect) measures that might include SETS to produce fairer results.

Not all teaching evaluations are problematic. Mid-semester course evaluations provide formative feedback to instructors and allow the opportunity for course corrections as needed to improve student learning. Using a learner-centered, small group instructional diagnosis (LC-SGID) approach, where the facilitated dialogue asks students to also focus on how their own actions reflect their learning in the course, researchers were able to demonstrate that such initiatives can actually result in students taking more responsibility for their learning than students who are asked only to reflect on the course or instructor (Hurney, Harris, Bates Prins, and Kruck, 2014). Yet most course evaluations at Indiana University occur at and are focused on the end of the semester only.

More troubling is the fact that a number of more recent studies (Mitchell and Martin, 2018; MacNeill, Driscoll and Hunt, 2015) have suggested that student evaluations of teaching advantage some faculty and disadvantage others (based on gender and race). They have shown that the language students use to describe male professors differs significantly from the language used to describe female professors. Online students rated instructors who they thought were men significantly higher than those they thought were female, even though the same instructor taught under both a male and a female identity. How such data is used by higher level administrators in personnel reviews may not be transparent to the faculty being reviewed. Reviewers often end up asking whether we should be using teaching evaluations at all in faculty personnel review.

Campuses across the US are taking note of this new research (Flaherty 2018). In response to these reports, the University of Southern California announced in Spring 2018 that they were doing away with the use of student evaluations of faculty teaching in tenure and promotion decisions. Students will still evaluate their professors but with a focus more on how engaged the student is in the class; faculty will be expected to explain how they used this student feedback to make course improvements. The data, however, will no longer be used as a direct measure of faculty performance. To replace evaluations, the university is investing in a new model of peer evaluation. The University of Oregon has also followed this same path and began to pilot a new Continuous Improvement and Evaluation of Teaching review system in Fall 2018 for similar reasons as the USC decision.

At Indiana University, two recent reports in response to President McRobbie’s 2017 State of the University charges on teaching highlight the need for renewed consideration of the role of SETs. *Identifying Pathways for Excellence in Teaching* (1 February 2019) points out that evidence for teaching excellence needs to be drawn from multiple sources focused on direct evidence of student learning. “Grades alone would not serve as direct evidence of student learning, nor would excellent (or poor) student end-of-course evaluations of teaching ...”. The report on

Assuring Pervasive Excellence in Teaching and Learning (May 2018) recommended that the university should “Explore more robust and precise models of student ratings of instruction that are designed to assess the learning that occurs in classes (IDEA, SALG).

The Task Force Process

The Taskforce met 7 times as a whole, with additional sub-committee meetings during the Fall 2019-Spring 2020 semester to carry out the following:

1. Review practices and history at Indiana University campuses.
2. Review research on teaching evaluation.
3. Explore nationally how campuses are responding to accumulating evidence for bias in student evaluations.
4. Discuss a new approach for Indiana University that will be:
 - Sensitive to accumulating evidence for SET bias
 - valid and reliable across different approaches to teaching and different disciplines.
 - feasible for departments/faculty to incorporate, and usable by administration across disciplines.

Members conducted interviews with officials at University of Oregon and University of Southern California regarding their transition, reviewed the recommendations for changes proposed at University of Massachusetts Amherst, reviewed IU’s current course evaluation items across campuses, and reviewed examples of processes and instruments used at multiple universities (including IU examples) for student feedback, teaching reflection, and peer review of teaching.

Recommendations

The committee identified the following best practices that seem to be emerging on the use of student feedback in assessing teaching excellence.

- 1) Due to the mounting evidence of bias, particularly gender and ethnic bias, in student ratings, eliminate the way that SETs are currently used for purposes of promotion and tenure and other personnel decisions. Faculty may choose to report results of student feedback for their courses in their dossier but they should not be required elements. Numerical results should be eliminated or minimized and should in no instance be used to compare one faculty member to another or to department or university averages.

We recommend that the UFC establish a formal policy for Indiana University in this regard so that numerical instructor ratings do not creep back into promotion and tenure and other personnel decisions.

NOTE: Policies needing possible review include ACA-27 (V-C); ACA-38 [Criteria for Promotion – Teaching]

- 2) Because SETs can provide valuable feedback to faculty on actual course instruction, course materials, particular pedagogies or teaching strategies employed in the course, we do not

recommend they be totally eliminated. They play an important role in formative evaluations of faculty teaching effectiveness. We recommend that IU reimagine SETS as follows: a short, standardized form that asks students to comment on two questions in a narrative form (no numerical ratings):

1. What worked in the class to support your learning?
2. What did not work in the class to support your learning?

Campuses, schools, departments and/or instructors would be able to add a few additional questions focused on student learning outcomes that are of interest, but evaluations should be kept brief, out of respect for the time we are asking students to invest to complete them. When possible, class time could be dedicated to completing feedback forms for this same reason. Following the collection of student feedback data, any hostile language should be redacted and campuses should provide expert consultation to instructors to help them interpret the feedback and implement changes.

Current practice has students completing these forms at the end of the semester. Research has shown, however, that mid-semester feedback opportunities provide important course correction information to instructors. Campuses should be encouraged to develop appropriate mid-semester student feedback opportunities and our tools should facilitate their use.

The form itself should be labeled “Student Experience Survey” in order to emphasize students’ responsibility for their own learning. Campuses and schools within campuses should be encouraged to define principles of excellent teaching which could then be used to guide their student feedback questions.

3) For purposes of promotion and tenure and other personnel decisions, Indiana University should develop two other processes to be used to provide evidence of faculty teaching effectiveness: annual faculty self-reflections on their teaching (see Appendix A) and a robust system of faculty peer review of teaching.

A. Annual Faculty Self-reflection Process: Reflective teaching is a self-assessment of teaching, wherein an instructor examines their pedagogy, articulates reasons and strengths for their strategies, and identifies areas for revision or improvement. We propose that each year, faculty write a self-assessment of the effect their teaching has had on their students' learning for each semester of teaching, and that this self-reflection becomes one of the multiple pieces of evidence used by faculty for the reappointment and promotion and tenure processes and other personnel reviews. Faculty members would be free to include student narratives drawn from SETs but would not be required to do so.

We recommend that each campus develop an instrument to guide faculty members through the process of creating a written teaching reflection. An example of such an instrument from the University of Southern California is included in Appendix A. This reflection would be attached to the faculty member’s annual report.

B. Peer Review Process: Peer review should provide both formative and summative feedback. In order to document satisfactory or excellent teaching for promotion and tenure, we recommend that faculty receive a peer review at least once every three years. For faculty claiming excellence in teaching, or who wish to emphasize development in teaching skills, more regular peer evaluations (e.g., one every one to two years) may be recommended. We also recommend that promoted faculty regularly receive a peer review at least once every three years, as a form of professional development. Faculty planning to apply for promotion (to Professor, Clinical Professor, or Teaching Professor) based on excellence in teaching may consider more regular peer evaluations (e.g., one every one to two years).

Each campus should develop a suitable process for peer review. Peer reviewers should either be compensated financially for their review efforts (just as Quality Matters peer reviewers are compensated), or peer review activities should count toward service expectations.

Peer reviewers should have been reviewed themselves and should ideally be experienced in peer review.

Peer reviewers should be faculty with an established record of excellence in teaching; this could include members of FACET, Trustees' Teaching Award winners, senior faculty who have been promoted on the basis of excellence in teaching, among other options. Faculty/staff of campus Centers for Teaching and Learning may also be good candidates as peer reviewers.

Finally, we recommend that peer reviewers complete FACET's peer review training course, available on Canvas to all faculty members.

References

Benton, Stephen L. and Cashin, William E. "Student Ratings of Teaching: A Summary of Research and Literature," *IDEA Paper #50*, 2010

Boring, Anne, Ottoboni, Kellie, and Stark, Philip B. "Student evaluations of teaching (mostly) do not measure teaching effectiveness," *ScienceOpen Research*, 7 January 2016.
<https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AETBZC.v1>.

Carpenter, Shana, Witherby, Amber E., and Tauber, Sarah K. "On Students' (Mis)judgments of Learning and Teaching Effectiveness," *Journal of Applied Research in Memory and Cognition*, 12 February 2020. <https://doi.org/10.1016/j.jarmac.2019.12.009>

Esarey, Justin and Valdes, Natalie, "Unbiased, reliable, and valid student evaluations can still be unfair," *Assessment & Evaluation in Higher Education*, 20 Feb 2020,
<https://doi.org/10.1080/02602938.2020.1724875>

Flaherty, Colleen. "Teaching Eval Shake-Up," *Inside Higher Education*, 22 May 2018.
<https://www.insidehighered.com/news/2018/05/22/most-institutions-say-they-value-teaching-how-they-assess-it-tells-different-story>

Holman, Mirya, Key, Ellen and Kreitzer, Rebecca. 2019 "Evidence of Bias in Standard Evaluations of Teaching." <https://docs.google.com/document/d/14JiF-fT--F3Qaefjv2jMRFRWUS8TaaT9JjbYke1fgxE/edit>

Hurney, Carol A., Harris, Nancy L., Bates Prins, Samantha C., and Kruck, S.E. "The Impact of a Learner-Centered, Mid-Semester Course Evaluation on Students," *Journal of Faculty Development*, 28(3): 55-62, 2014.

MacNell, Lillian, Driscoll, Adam, and Hunt, Andrea N. "What's in a Name: Exposing Gender Bias in Student Ratings of Teaching," *Innovative Higher Education*, 40(4): 291-303, 2015.

Mitchell, Kristina M.W. and Martin, Jonathan. "Gender Bias in Student Evaluations," *PS: Political Science & Politics*, 51(3): 648-652, 2018.

Stark, Philip B. and Freishtat, Richard. "An Evaluation of Course Evaluations," *ScienceOpen*, 2014. Available at <https://www.stat.berkeley.edu/~stark/Preprints/evaluations14.pdf>

Uttl, Bob, White, Carmela A., and Gonzalez, Daniela Wong. "Meta-Analysis of Faculty's Teaching Effectiveness: Student evaluation of teaching ratings and student learning are not related," *Studies in Educational Evaluation*, 54 (September): 22-43, 2017.

Vasey, Craig and Carroll, Linda. "How Do We Evaluate Teaching?," *Academe*, 102(3): May-June 2016. <https://www.aaup.org/article/how-do-we-evaluate-teaching#.WrFBzOgbNPY%0A>

Informational Interviews:

Ginger Clark, University of Southern California
 Sierra Dawson, Associate Vice Provost for Academic Affairs, University of Oregon

Reports:

Recommendations for Changes to the Evaluation of Teaching, University of Massachusetts Amherst, Fall 2017

Evaluations Focused on Student Learning Outcomes:

IDEA Student Ratings of Instruction: <https://www.ideaedu.org/Services/Services-to-Improve-Teaching-and-Learning/Student-Ratings-of-Instruction>

Student Assessment of their Learning Gains (SALG): <https://salgsite.net>

Appendix A

Faculty Self-Reflection on Teaching

1) Instructional Goals for the year: Discuss your measurable goals as an instructor from the previous year (examples: I wanted to flip my class, change participation to in-class work, increase active learning, etc.)

2) Instructional Practices: Discuss specific examples of your teaching practices that support your goals identified above. Examples might include:

- Active-learning strategies
 - Give specific examples of, and discuss the value of, activities and strategies you use in courses to help students engage with and process course content
- Key assignments/assessments
 - Give specific examples of, and discuss the value of, assignments and assessments you assign students in order to accurately evaluate their mastery of your course learning objectives. For each, also specify the corresponding course learning objectives.
- Course materials
 - Give specific examples of, and discuss the value of, course materials (such as readings, videos, podcasts, etc.) you use in your courses.

3) Outcomes of Instructional Goals for the year: Discuss the outcomes of your efforts. Present clear, concise evidence of whether your goals were accomplished, and lessons learned.

4) Instructional Alignment: Discuss how your instructional practices align with department, school, and University goals for teaching. Each campus faculty should frame and adopt clear statements of what these instructional goals are for their campus and its mission and students.

5) Areas for Improvement: Articulate specific areas for improvement or changes to teaching practices based on student outcomes, student course feedback, or other data.

6) Improvement Goals for next year: Delineate at least two (2) clear, concise, measurable goals for improving your teaching practices for the next academic year, aligned with the above areas for improvement.