

Joint Task Force on Teaching Assessment and Evaluation Recommendations April 23, 2019

Sponsors:

**Peter Martin, Faculty Senate
Ann Marie Van Der Zanden, Provost's Office**

Co-Chairs:

Jonathan Sturm [Senate/LAS-MUS], Sara Marcketti [CHS/CELT]

**Jan Lauren Boyles [GSJC/COMM]
David Cantor [BUS/SCIS]
Amanda Fales-Williams [VET/PATH]
Rose Martin [HS/FSHN]
Dave Peterson [LAS/POL SCI]
Jo Anne Powell-Coffman [CALC/GDCB]
Mikesch Muecke [DES/ARCH]
Diane Rover [ENG/ECE]
Henry Schenck [LAS/MATH]
Norin Yasin Chaudhry [GPSS]
Zahra Barkley [SG]**



Joint Task Force on Teaching Assessment and Evaluation

Recommendations

April 23, 2019

Summary: Teaching is at the core of Iowa State University's mission, and excellent teaching is a practice its faculty strive to achieve and maintain. Evaluating teaching is a critical process toward establishing and acknowledging excellence, and as such is worthy of periodic reviews and updates. The Joint Task Force on Teaching Assessment and Evaluation was convened in August 2018 by the Faculty Senate and the Office of the Senior Vice President and Provost with the charge to: (1) Review existing best practices regarding holistic teaching assessment and evaluation; (2) Review procedures and processes by which teaching is assessed and evaluated on campus; and (3) Provide recommendations for how the assessment, evaluation and promotion of teaching could be reviewed at Iowa State University (ISU). As part of its work, the Task Force consulted with external experts who study best practices in teaching assessment, studied universities across the nation that are also reconsidering their approaches to teaching evaluation, and reviewed teaching evaluation processes on ISU's campus.

In recent years, the limitations of student end-of-semester course feedback have been studied intensively. Current best practices in teaching assessment indicate that, when evaluating an instructor's effectiveness and impact upon the university's teaching mission for decisions regarding promotion, advancement, reward, or merely for reflection toward an improved teaching practice, multiple sources of information need to be considered in order to build a complete picture. Further, the processes that integrate measures of teaching effectiveness should be transparent and communicated broadly.

In this report, the Joint Task Force on Teaching Assessment and Evaluation summarizes its yearlong activity and provides recommendations for strengthening the Evaluation of Teaching process at ISU across three broad areas of emphasis outlined in phases for implementation. The recommendations consider changes to student course evaluations, changes to the Faculty Handbook in support of best practices, and the development of rubrics to guide, clearly describe, and strengthen the process of evaluating teaching effectiveness.

A. Background: The land-grant act established colleges and universities with a three-part mission encompassing, teaching, research, and extension/outreach. At ISU, teaching is an essential aspect of our mission to "Create, share, and apply knowledge to make Iowa and the world a better place." ISU's strategic vision, 2017-2022 makes clear that the first goal of our institution is to "Ensure access to the ISU Experience – including an exceptional education offering practical, global, and leadership experiences that shape the well-rounded citizens and informed critical thinkers needed in the 21st century."

In addition to its vital role in the mission and strategic vision for the university, teaching is an essential component of the overall evaluation process for faculty with teaching responsibilities (Faculty Handbook 5.1.1.2). Effective teaching is a part of personnel decisions made during hiring, annual reviews and contract renewals, and is central to promotion, tenure and to advancement decisions for term faculty. Because these

applications are so important to the faculty and the university, evaluation instruments and processes must be equitable, transparent, and should manifest the best practices currently understood across the academic environment. While the ISU Faculty Handbook (5.2.2.3.1) emphasizes that documentation should include “materials such as teaching philosophy, student ratings of teaching, teaching materials and forms of assessment, peer evaluations based on both classroom observations and review of teaching materials, and evidence of student learning,” in practice student evaluations of teaching have assumed a primary role in determining teaching effectiveness at ISU, as well as at other universities across the nation. Peer-reviewed research and empirical evidence indicates that the primary emphasis on student evaluations of teaching as a measure of teaching effectiveness is an incomplete and flawed measure and one with significant potential problems for faculty, students, and the institution at large. The Task Force’s assessment of our current practices and our recommendations for the future are a first step toward a practical reconsideration of the evaluation of effective teaching at ISU towards a more holistic process leading to more detailed and nuanced results.

Task Force Membership and Process

Associate Provost Ann Marie VanDerZanden and Faculty Senate President Peter Martin are the Executive Sponsors of the Task Force.

Following two years of preliminary work by Faculty Senate committees, the current Task Force on Teaching Assessment and Evaluation was formed in August 2018. Co-chaired by the Faculty Senate and the Center for Excellence in Learning and Teaching (CELT), the Task Force includes representatives from across the university, including student representatives and faculty from diverse colleges and disciplines. The membership included:

- Jonathan Sturm, Faculty Senate President-Elect, Professor, College of Liberal Arts and Sciences, (co-chair)
- Sara Marcketti, CELT Director, Professor, College of Human Sciences (co-chair)
- Zahra Barkley, Student Government
- Jan Lauren Boyles, Assistant Professor, Greenlee School of Journalism
- David Cantor, Professor, Ivy College of Business
- Norin Chaudhry, Graduate and Professional Student Senate
- Amanda Fales-Williams, Tyrone D. Artz, M.D. Chair for Teaching Excellence, College of Veterinary Medicine
- Rose Martin, Senior Lecturer, College of Human Sciences
- Michael Muecke, Associate Professor, College of Design
- David Peterson, Professor, College of Liberal Arts and Sciences
- Jo Anne Powell-Coffman, Professor and former Department Chair, College of Agriculture and Life Sciences
- Diane Rover, University Professor, College of Engineering
- Henry Schenck, Department Chair and Professor, College of Liberal Arts and Sciences

The Task Force met monthly throughout the 2018-19 academic year. Our charge was to review existing best practices regarding holistic teaching assessment and evaluation, review procedures and processes by which teaching is assessed and evaluated on campus, and provide recommendations for how the assessment, evaluation and promotion of teaching could be improved at Iowa State University.

Support from an Association of American Universities (AAU) STEM Network mini-grant entitled “New Heights through New Networks” enabled the Task Force to begin in September 2018 with two visits by nationally recognized experts in this field: Elizabeth Barre from Wake Forest University and Andrea Greenhoot from the University of Kansas. Work continued throughout the fall as the Task Force worked in smaller teams and researched other American universities that also are studying teaching evaluation. We discovered how they addressed current research and what new policies or practices they were considering or implementing.

In the winter and early spring, the Task Force again worked in sub-groups to consolidate our prior external research and to draft recommendations in the areas of 1) the student ratings survey tool, 2) suggested new Faculty Handbook language, and 3) rubrics and their use. Our final full-group meetings discussed the draft recommendations and finalized this report.

This report provides a summary of the Task Force deliberations as well as providing a foundation to move forward. Our deliberations were informed by extensive study of the existing research on the topic of teaching documentation, evaluation, and assessment. Based upon our research and discussions, our principal goals include:

- moving the ISU community meaningfully toward viewing effective teaching and its evaluation as a formative tool to improve teaching at ISU,
- moving our evaluation process meaningfully toward a more holistic process in assessing faculty teaching, and
- de-emphasizing the past use of student end-of-semester course ratings as a unilateral means for comparing faculty performance within departments and across colleges or as a unilateral data point in the promotion, advancement and recognition of faculty.

B. Review of best practices nationwide and of current ISU practices

There is increasing national attention to best practices for assessing and encouraging effective teaching. Teaching assessments play critical roles in the continual improvement of courses and curricula, and they are central to faculty recognition and reward structures. Current research indicates that while student end-of-semester evaluations are one important data point in a complete evaluation of teaching, they should not be an exclusive data point.

Historically, universities have relied on end-of-semester Lickert scale student surveys as the primary indicator of good teaching. This approach is now understood to be flawed: Student course ratings are influenced by biases, and the ratings are not strongly correlated with student learning gains (Clayson et al., 2009, MacNell et al., 2015; Linse 2017, Uttil

et al., 2017; Mengel et al., 2019). Current best practices leverage multiple sources of information when assessing effectiveness of an instructor in a course and when evaluating a faculty member's summative contributions to the teaching mission (Finkelstein and Keating, 2018).

At ISU, it is widely acknowledged that information beyond student course evaluations can be valuable, but there is still an over reliance on end-of-semester student surveys. We do not yet have clear guidelines about how to weigh and integrate multiple sources of information when evaluating teaching effectiveness.

Current process at ISU: strengths and weaknesses of practices that emphasize student ratings

While the ISU Faculty Handbook provides several examples for ways to document teaching effort, for the majority of personnel decisions and teaching award considerations, student evaluations of teaching have become the primary means of evaluation. For the majority of academic departments, two questions on the student evaluations of teaching guide personnel decisions: "Overall instructor" and "Overall course" ratings. After evaluating current research and studying other universities' decisions regarding evaluating teaching, the Joint Task Force sees the following strengths and weaknesses with the current evaluation process and emphasis at ISU:

Strengths:

- Current student evaluations of teaching provide student feedback to instructors on their experiences in courses.
- The electronic system of conducting student ratings of teaching is relatively quick and simple to complete.

Weaknesses:

Based upon a literature review, external expert opinions, and perspectives from faculty and students, there appear to be a number of weaknesses in the current process of teaching evaluation broadly conceived as implemented at ISU. Principally:

- Ratings surveys are currently made open at random times in the semester depending upon which departments are running them. Students do not appreciate the constant email reminders about ratings still needing to be completed, and faculty have lost control of the timing of their course's evaluation.
- The results of end-of-semester ratings are not distributed to all faculty in a consistent and timely manner. In some departments, faculty do not receive prior semester feedback until 2-4 months following the conclusion of a semester, which inhibits faculty using the feedback to adjust their teaching in a subsequent semester.
- Student evaluations, as currently employed, feature too prominently in annual evaluation decisions, P&T reviews, and faculty awards. Chairs and committees often rely upon them either entirely or too much for quick

assessments of faculty teaching, thus missing the nuance and diversity of faculty work in the area of teaching.

- Questions in surveys as currently employed allow too much opportunity for biases to be a factor in student answers, which creates an uneven and unfair environment for females and people from diverse racial/ethnic backgrounds in their employment and career trajectories.
- Questions in many current evaluations across the university focus too much on summative perceptions and overlook formative input that would actually help faculty improve their teaching.
- Questions do not consistently ask students across the university to self-evaluate or assess their contribution to their own learning, which is a contributing factor to their success or failure in a course.
- Current surveys lack contextual/grade distribution data. Grade distribution data (even course measures such as mean grade in the class, and the historical 5 year mean) are crucial to putting student evaluations in the proper context. A student evaluation average of 5.0 loses its luster when the mean student grade in the class is 4.0. Contrarily, multiple studies indicate that rigorous grading can lower student evaluation scores.
- There are no questions outside of the Overall Instructor and Overall Course questions that are universally used across the university to create a sense of consistency and a baseline for university-wide assessment. The Task Force feels these are not strong measures of teaching effectiveness. There is concern that, by dismissing any additional evaluation questions for promotion or tenure decisions, the university is missing valuable data points in a faculty teaching profile.

Key desirable characteristics of student ratings

The Task Force sees the most desirable aspect of student ratings being their ability to inform instructors about student perceptions of their teaching—both the successful attributes and areas needing improvement at key points in a semester.

The Task Force values the ideal that, through a carefully considered and implemented process, students can be offered more than just an end-of-semester opportunity to provide input on a course. An early formative survey that catches concerns early, allowing an instructor to make corrections to a course or address publicly why student suggestions may or may not be possible, is likely to improve instructor-student trust and collaboration toward a successful semester for all involved parties. Such an early formative survey would be made available exclusively to the instructor and would not serve any summative evaluation role.

Additional data points for inclusion in a holistic review of teaching

Having reviewed current literature and having met with scholars who study these issues, the Task Force accepts that the holistic review of teaching should include many sources of input. These may include course materials, faculty self-reflection, peer review of teaching, external review (as appropriate), and other aspects of the complete teaching

work profile unique to individual faculty as well as a thoughtful student survey that focuses upon formative reflection. A rubric that meets departmental, college, and university needs and expectations and that visibly encapsulates the wide spectrum of input on an instructor's teaching would make assessment more clear and fair.

Other institutions, organizations and scholars have approached these issues and have piloted new strategies to improve evaluations of teaching effectiveness. The Association of American Universities (AAU) has encouraged leading research universities to be proactive in aligning institutional reward systems with best practices (Dennin et al, 2017). The Task Force met with National Science Foundation grant recipient Dr. Andrea Greenhoot, the University of Kansas's Director of the Center for Teaching Excellence and Professor of Psychology, who is helping departments at the University of Kansas to develop rubrics for evaluations of teaching (<https://www.aau.edu/sites/default/files/Blind-Links/KU-Benchmarks-Framework.pdf>). The University of Massachusetts at Amherst and the University of Colorado at Boulder are also drawing on these materials to scaffold more holistic evaluations of teaching on their campuses. ISU can build on these NSF funded materials developed and lessons learned from peer institutions.

C. Task Force Recommendations

The Task Force proposes recommendations concerning three areas:

- The Student end-of-semester ratings tool
- Faculty Handbook language and procedural guidance
- Rubrics: their content, and their use.

The implementation phases suggest a 2-4 year process resulting in sequential steps toward an improved model for evaluating teaching at Iowa State University. For example, Phase 1 would be an initial step in the process.

Recommendation 1: Student feedback and the end-of-semester ratings tool

Phase 1

- Change ISU current designations of end-of-semester survey tools from Student Evaluations of Teaching (SET) or Student Evaluations of Instruction (SEI) to Student Ratings of Teaching to align with recent changes in the literature to describe these tools. (See also Recommendation 2, Phase 1.)
- Allow for class time (in courses that meet face-to-face) for students to complete the student ratings.
- Provide faculty with access to student ratings immediately following end-of-course grade submission (for classes larger than 20-25 students where anonymity is possible). End-of-semester ratings should be distributed to all faculty within two weeks of the end of a semester.
- Include a statement at the beginning of the survey that explains that student ratings are important as an educational tool for improving teaching at ISU.

- Include a separate statement that outlines research on existing biases related to student ratings and asks students to be mindful of these issues in their responses.
- If incentives are considered as a means to increase student participation in end-of-semester ratings, consider orienting them toward increased percentages of participation as a whole class not individual participation. For example, incentives could be offered to an entire class of students if student participation in survey ratings reached 85% or higher.

Phase 2

- Consider other platforms and approaches for administering end-of-semester student ratings besides Class Climate.
- Initiate a process to create a pool of formative survey questions.
 - Consider eliminating the “Overall Professor” and “Overall Course” questions and replace them with more student-reflective questions. Emphasize formative over summative feedback.
 - Standardize several questions on the end-of-semester surveys across the entire campus and all departments for better consistency.
 - Implement a formative survey early in the semester. The results would only be seen by the instructor teaching the course/section as an early feedback tool.
- Provide department standard deviations along with department means for Likert scale questions to indicate the spread of ratings within a department.

Phase 3

- Consider implementing an exit survey at graduation that asks students to identify the top 2-4 courses in their major and the top 2-4 courses outside their major that influenced them positively (made the longest-lasting impression upon them) while students at ISU. (This provides a separate opportunity for students to revisit their curriculum after their initial course impressions have possibly changed and with the perspective that time, distance, and perhaps maturity can provide.)

Recommendation 2: Language in the Faculty Handbook and guidance for the implementation and use of student ratings

Phase 1

- Rename within the Faculty Handbook and in college and department Faculty Handbooks/Governance Documents the term “Student Evaluations of Teaching” to “Student Ratings” to delineate clearly that these are drawn from student opinions. Students typically have neither the content knowledge, nor the training to truly evaluate teaching.
- Define the evaluation of teaching as a holistic perspective on teaching deriving from multiple qualitative and quantitative points of information, including: student ratings, peer evaluation, faculty self-evaluation and reflection, course materials, advising and mentorship, curriculum development.

- The Faculty Senate should require through the Faculty Handbook that all departments have a rubric for evaluating the diverse aspects of teaching (for example teaching practices, course goals, content and alignment, achieving learning outcomes, classroom climate, reflection and iterative growth, mentoring/advising, involvement in teaching service and community) that both meets department needs and is approved by their college. Provide a date for meeting this expectation, perhaps by 2022.
- Stipulate that neither a “below expectations” annual evaluation nor a faculty teaching award can be based exclusively upon student ratings numbers.

Phase 2

- Stipulate that, if university-wide questions are adopted as a standard, departments must include them in their individual student surveys.
- Stipulate that every department shall create language in their governance document that defines the process and content of teaching evaluation for the department and that is approved by their college.

Recommendation 3: Develop rubrics to be used by departments in the holistic evaluation of teaching either in tandem with or independently of other recommendations

Phase 1

- Identify departments that wish to be early adopters of teaching evaluation rubrics. Provide the opportunity for early adopters to create their own rubric or use an existing and respected rubric. Provide incentives and opportunities for early adopters to share their experiments and findings with other campus units.
- Provide workshops through CELT on creating teaching evaluation rubrics.
- Encourage departmental retreats to discuss and develop teaching effectiveness paradigms in alignment with this Task Force’s recommendations. Provide incentives with remaining funds from the AAU mini grant in 2019.

Phase 2

- Require all departments to adopt or create a teaching evaluation rubric that meets their unique needs by 2022. The rubric will assess teaching effectiveness broadly and inclusively as understood by each discipline, but will also be approved by the college in line with the Faculty Handbook timeframe mentioned above.
- Encourage departments to develop more regular schedules for evaluating teaching across multiple sources of information, as well as to clarify language that ensures all faculty understand this process, particularly its timing and intent.

Additional recommendations

- Create more opportunities university-wide to reward teaching excellence.
 - Encourage departments to develop department-level teaching awards.
- Encourage a university-wide conversation on what effective teaching is (defined) versus ineffective teaching (defined).
- Increase instruction and training for chairs to enhance their expertise in collecting and evaluating all data submitted for teaching evaluation.
- Consider implementing rotating departmental CELT ambassadors (along the lines of Library liaisons) so that, over a period of time, all faculty interact with CELT.
- Recognize and credit faculty efforts at self-improvement in their teaching through internal (CELT) or external workshops.
- Credit faculty actions (workshops) to increase inclusivity in teaching environments.
- Consider including the *average* grade distribution of all students in a course or section to correlate student ratings with the potential for grade inflation.

References:

- Clayson DE (2009) Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education*. 2009;31(1):16–30
- Dennin M et al (2017) Aligning practice to policies: Changing the culture to recognize and reward teaching at research universities. *CBE Life Science Education* 16: [10.1187/cbe.17-02-0032](https://doi.org/10.1187/cbe.17-02-0032)
- Finkelstein N and Keating J (2018) Promoting scholarly evaluation of teaching: Addressing the third rail of academia. *Science News* Aug 13, 2018. <https://www.ascb.org/science-news/promoting-scholarly-evaluation-teaching-addressing-third-rail-academia/>
- Linse AR (2017) Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation*. 54: 94-106
- MacNeill L, Driscoll A, and Hunt AN (2015) What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*. 40: 291-303
- Mengel, F, Sauermann J, and Zolitz U. (2019) Gender Bias in Teaching Evaluations. *Journal of the European Economic Association* 17 (2), 535-566
- Uttl, B., White, CA and Gonzalez DW (2017), [Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related](#), *Studies in Educational Evaluation*, 54: 22-42.