

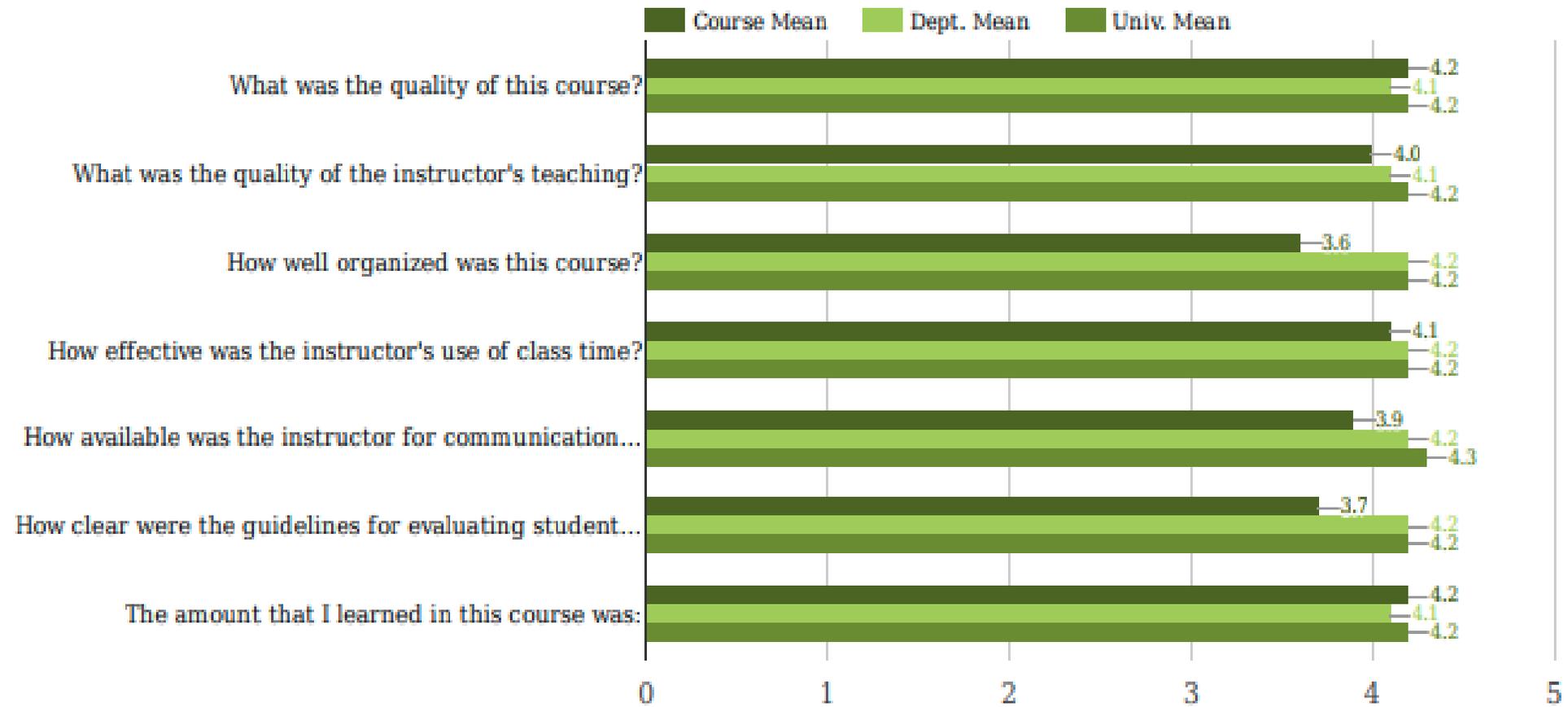
# Improving how we evaluate teaching

Senate Meeting Jan 31, 2018

UO Senate Teaching Evaluation Task Force  
Co-chairs Sierra Dawson, Bill Harbaugh  
University of Oregon

- Evaluation of teaching has 2 main purposes:
  - Continual Improvement
  - Evaluation for contract renewal, promotion, merit, awards.
- UO and most universities currently rely on the same Student Evaluations of Teaching (SET's) for **both** these purposes (some use of peer evaluations etc).
- SETs are Primarily numerical scores of student's answers to Likert scale questions. Students can also give written comments, which are only used for assessment if signed.
- Collected by computer outside of class, last week of course, before grades are released. High response rate because students can see their grades immediately after completing, otherwise must wait.

Please share with us your basic perceptions of the course.



# Use of Student Evaluation of Teaching (SETs) Scores

- Used by faculty for info on how to improve their teaching
- Numerical ratings were the first screen for teaching awards
- For departments, the scores are the main input into evaluating teaching for merit raises, promotion etc.
- Scores from other schools are used for our faculty hiring decisions
- SETs are used by our PhD's going on job market.

# Some problems with using SETs:

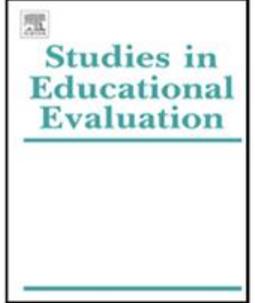
- Questions are not specific enough to provide useful input for improving teaching
- Students rush through them (up to 8/term?)
- Scores are biased by gender, race/ethnicity, international background etc.
- Scores are very highly correlated with grades – incentive for grade inflation
- Faculty who adopt innovative, evidence based teaching methods that require higher student effort can get lower scores, particularly at first
- At UO, **negative** correlation between scores and learning outcomes were found - similar to National meta-analysis



Contents lists available at [ScienceDirect](#)

## Studies in Educational Evaluation

journal homepage: [www.elsevier.com/stueduc](http://www.elsevier.com/stueduc)



# Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related

Bob Uttl\*, Carmela A. White<sup>1</sup>, Daniela Wong Gonzalez<sup>2</sup>

*Department of Psychology, Mount Royal University, Canada*

selection bias. Despite more than 75 years of sustained effort, there is presently no evidence supporting the widespread belief that students learn more from professors who receive higher SET ratings. **If anything, the latest large sample studies show that students who were taught by highly rated professors in prerequisites perform more poorly in follow up courses** (Weinberg, Hashimoto, & Fleisher, 2009; Yunker & Yunker, 2003).

In turn, our findings indicate that depending on their institutional focus, universities and colleges may need to give appropriate weight to SET ratings when evaluating their professors. Universities and colleges focused on student learning may need to give minimal or no weight to SET ratings. In contrast, universities and colleges focused on students' perceptions or satisfaction rather than learning may want to evaluate their faculty's teaching using primarily or exclusively SET ratings, emphasize to their faculty members the need to obtain as high SET ratings as possible (i.e., preferably the perfect ratings), and systematically terminate those faculty members who do not meet the standards. For example, they may need to terminate all faculty members who do not exceed the average SET ratings of the department or the university, the standard of satisfactory teaching used in some departments and universities today despite common sense objections that not every faculty member can be above the average.

# So why do we continue to use this failed system?

- Students want to have a voice.
- Dept heads need early warning of serious problems.
- Administrators need some information for hiring and promotion.
- For merit raises departments must rate faculty – even if this rating is not consistent with the university's academic mission
- Inertia – it is hard to change a system

# AAU is calling for a new system:

## Aligning Practice to Policies: Changing the Culture to Recognize and Reward Teaching at Research Universities

Michael Dennin,<sup>†</sup> Zachary D. Schultz,<sup>‡</sup> Andrew Feig,<sup>§</sup> **Noah Finkelstein**,<sup>||</sup> Andrea Follmer Greenhoot,<sup>¶</sup> Michael Hildreth,<sup>#</sup> Adam K. Leibovich,<sup>@</sup> James D. Martin,<sup>\*\*</sup> Mark B. Moldwin,<sup>††</sup> Diane K. O'Dowd,<sup>‡‡</sup> Lynmarie A. Posey,<sup>§§</sup> Tobin L. Smith, and **Emily R. Miller**

Recent calls for improvement in undergraduate education within STEM (science, technology, engineering, and mathematics) disciplines are hampered by the methods used to evaluate teaching effectiveness. Faculty members at research universities are commonly assessed and promoted mainly on the basis of research success. To improve the quality of undergraduate teaching across all disciplines, not only STEM fields, requires creating an environment wherein continuous improvement of teaching is valued, assessed, and re-warded at various stages of a faculty member's career. This requires consistent application of policies that reflect well-established best practices for evaluating teaching at the de-partment, college, and university levels. **Evidence shows most teaching evaluation practices do not reflect stated policies, even when the policies specifically espouse teaching as a value. Thus, alignment of practice to policy is a major barrier to establishing a culture in which teaching is valued.** Situated in the context of current national efforts to improve undergraduate STEM education, including the Association of American Universities Undergraduate STEM Education Initiative, this essay discusses four guiding principles for aligning practice with stated priorities in formal policies: 1) enhancing the role of deans and chairs; 2) effectively using the hiring process; 3) improving communication; and 4) improving the understanding of teaching as a scholarly activity. In addition, three specific examples of efforts to improve the practice of evaluating teaching are presented as examples: 1) Three Bucket Model of merit review at the University of California, Irvine; (2) Evaluation of Teaching Rubric, University of Kansas; and (3) Teaching Quality Framework, University of Colorado, Boulder. These examples provide flexible criteria to holistically evaluate and improve the quality of teaching across the diverse institutions comprising modern higher education.

# What has the Teaching Evaluation task force been doing?

- Meeting every two weeks since spring term
- Talking with other Universities tackling the same problem
- Meeting with faculty, graduate students and undergraduate students
- Developing and modifying prototypes for Student Experience Surveys to replace current course evaluations
- Working towards a framework of Teaching Evaluation for contract renewal, merit, promotion and tenure that is not focused on course evaluation ratings

# Task Force Members

**Helen Chu:** Associate Dean of Libraries

**Edward Davis:** Curator, Museum of Natural and Cultural History, Assistant Professor Earth Sciences

**Sierra Dawson:** Assistant VP, Provost and Academic Affairs, Human Physiology Faculty

**Judith Eisen:** Professor, Biology

**Nina Fox:** LMS Administrator & Manager, UO Libraries

**Andy Halvorsen:** Career Instructor, American English Institute

**Bill Harbaugh:** Professor, Economics and Senate VPres

**Austin Hocker:** Student, Human Physiology

**Brian Lowery:** Registrar's Office

**Lee Rumbarger:** Director Teaching Engagement Program

**Christopher Sinclair:** Associate Professor, Mathematics and Senate Pres

**Sanjay Srivastava:** Associate Professor, Psychology

**Emily Wu:** Student, Economics

# Who have we met with?

Associate Deans

Grad Council

ASUO Exec

SWAT

Graduate Student Advisory Board

Women's Center

Dean of Student Life - staff

Student Trustee/Senators

Women of Color Coalition

Mujeres

Business Dean's Student Advisory  
Council

Dean of Students Advisory  
Committee

IMPACT -Intercultural Mentoring

Program Advancing Community Ties

# New system for evaluation:

Departments and faculty have three needs:

- 1) Quickly identify and correct serious concerns.
- 2) Evaluate the quality of teaching (for promotion, merit).
- 3) Give faculty information and incentives to improve teaching and learning.

# Principles:

- Ensure all three voices are included in Teaching Evaluation (student, faculty peers, self).
- Ensure the student's role in the process provides the opportunity to have a voice, protects their anonymity, is efficient, does not rely on negative incentives, and cannot be used as stand-alone Teaching Evaluation.
- Mitigate bias at each step of the teaching evaluation process.

# How we propose to move forward:

- Continue to talk to stakeholders: Students, Senate, OPAA, Deans, Dept heads, faculty etc.
- Finalize drafts of Student Experience Surveys (midterm and end of term) and continue to test them.
- Bring new holistic Teaching Evaluation policy to the Senate by Spring.