



Academic Senate

AD HOC COMMITTEE ON EVALUATION OF TEACHING

January 20, 2021

To: Jason Stajich, Chair
Riverside Division

From: David Nelson, Chair 
Ad Hoc Committee on Evaluation of Teaching

Re: Final Report of the Ad Hoc Committee on Evaluation of Faculty Teaching

Dear Senate Colleagues:

The Ad Hoc Committee on Evaluation of Teaching was convened in February 2018. The committee was charged with reviewing the current policies, procedures, and mechanisms for how UCR faculty members are evaluated for their teaching and was tasked with the following:

- 1. Establish policies and procedures for accurate and reliable evaluation of faculty teaching in undergraduate, graduate, and professional courses. Recommendations should take into account the Academic Personnel requirements outlined in APM 210-1-d.1 that requires more than one kind of evidence of teaching evaluation in each faculty review file.*
- 2. Review iEval, used as the primary student evaluation instrument of faculty members of teaching, for its efficacy in all courses. If appropriate, recommend changes that maximize student participation and provide faculty members with constructive feedback to improve their teaching. Any recommendations on evaluation instruments should take into account appropriateness of course level as well as differences in pedagogical approaches (i.e., traditional lectures, online offerings, hybrid classes, etc.).*

In 2020, in coordination with the Director of Evaluation and Assessment, Omar Safie, we conducted a survey of the experiences and opinions of UCR faculty with the current iEval system. This survey indicated widespread dissatisfaction with the current system and its use as a measure for teaching effectiveness (see Appendices B-E of Subcommittee A report).

Here we present the recommendations that we have developed in response to the Senate's charges, having considered the concerns of UCR faculty and well-documented problems with student evaluations of teaching. Our recommendations are detailed in the attached summary reports from two subcommittees. Subcommittee A was tasked with proposing revisions to iEval that could make it a more fair and effective instrument. Subcommittee B was tasked with providing recommendations of how the campus should interpret and responsibly use teaching evaluations for merit and promotion decisions.

While our work was underway, similar reviews of problems and best practices for student evaluations of teaching were undertaken by the Academic Council Teaching Evaluation Task Force and the UC Centers for Teaching and Learning. As detailed in our last communication with the Senate, we have taken their reports under consideration and note that in general we have come to similar conclusions. We highlight several particularly relevant recommendations from the 2019 report from the UC Centers for Teaching and Learning report in Appendix A of the Subcommittee A report.

On the following page we propose a two-phased approach to improve our current system of teaching evaluation that draws from our Subcommittees' reports. Phase I consists of several stopgap measures that we anticipate could be implemented as soon as this current winter quarter. Phase I recommendations are not intended to be long-term solutions in themselves, and Phase II will be required to bring about meaningful change.

Sincerely,
The Ad Hoc Committee on Evaluation of Teaching

Chair, David C Nelson, Botany and Plant Sciences
Michelle E. Bloom, Comparative Literature and Languages
Kevin Costello, Mathematics
Long Gao, School of Business
Jean Helwege, School of Business
John M. Heraty, Entomology
Jennifer Hughes, History
John Levin, Graduate School of Education
Suveen Mathaudhu, Mechanical Engineering
John Stamp, University Writing Program
Ex Officio Jennifer Lynn Brown, Undergraduate Education (Vice Provost for Undergraduate Education)
Ex Officio Richard A. Cardullo, Evolution, Ecology, and Organismal Biology
Ex Officio Daniel Jeske, Statistics (Vice Provost of Academic Personnel)
Ex Officio Mariam B. Lam, Comparative Literature and Languages (Vice Chancellor of Diversity, Equity, and Inclusion)
Ex Officio Josh Bright, Interim Associate Vice Chancellor Chief Information Officer

Phase I

- Add a preamble about bias in student evaluations of teaching (SETs) at the beginning of the iEval form.
- Replace iEval question #5 (“I normally spent at least two hours preparing for each hour of class”), which was considered one of the least informative iEval questions by the committee, with “I have considered the possibility that personal biases around race, gender identity, age and sexual orientation may impact my responses to this survey.”
- Stop including department and campus comparative statistics (percentiles, mean, median, SD) in the numerical iEval report.
- Communicate to UCR faculty and P&T evaluating bodies the importance of additional forms of evidence of teaching effectiveness that supplement the iEval SET. This should include a reminder of the requirement in APM 210-1-d.1 that “More than one kind of evidence [of teaching effectiveness] shall accompany each review file.” Although “opinions of students” is only one of the significant types of evidence listed, it dominates our current P&T evaluation process. Specify different forms of evidence of teaching effectiveness that will be considered acceptable by evaluating bodies.
- In coordination with the previous recommendation, modify eFile so that faculty can easily include different types of evidence of teaching effectiveness. That is, add new categories of evidence to eFile that can be selected and uploaded. The current designation of “Other Teaching Info” implies that alternative forms of evidence, which must further be categorized as “Other” Activities, do not carry as much weight as SETs listed under Teaching Records. More explicit instructions on what constitutes useful information in a self-reflective Teaching Statement could also be provided.

Phase II

The next phase will entail the re-design of the current student evaluation of teaching to produce a more equitable and useful tool for evaluation of teaching effectiveness and pedagogical improvement. In the attached reports from Subcommittees A and B, we have made a set of recommendations that provide guidance on issues that should be addressed in an iEval revision and in the continued use of student evaluations of teaching for promotion and tenure decisions. These recommendations are largely congruent with the 2019 UC Centers for Teaching and Learning Report, which should also be taken into consideration. Although we have provided examples of how the evaluation could be modified, this committee recommends that the exact wording and selection of questions should be designed by experts in pedagogy based upon current research.

SUBCOMMITTEE A (revision of iEval student evaluation of teaching)

Charge

Propose revisions to iEval to become “*an instrument that is more closely tied to student learning, rather than just satisfaction, and that is tailored to the specific class and its format, teaching practices, and learning goals. ... to develop a course evaluation that provides better information to faculty, encourage students to reflect more thoughtfully about their educational experiences and enhances the learning partnership*” (CEC report, Stanford 2013)

Background and Process

The format of UCR’s Student Evaluation of Teaching has been unchanged for almost fifteen years, since 2006, with the exception that in the fall of 2009 UCR Academic Senate approved that assessment would be moved online. The Senate formed an Ad Hoc committee to examine Student Evaluation of Teaching in 2015-16. That committee’s recommendations were rejected by Senate Academic Council in Winter of 2016 as not being far-reaching or substantive enough in their proposed revision. A new Ad Hoc committee was formed in the spring of 2018 and has been meeting ever since. In 2019 the committee was divided into two subcommittees, with Subcommittee A dedicated to reviewing UCR’s iEval instrument with a mandate to make suggestions so that it might become more closely tied to student learning.

Subcommittee A compared teaching evaluations from multiple campuses, studied the 2020 UC Teaching and Learning Centers Report (UCTLCR), and conducted a survey of faculty for opinions on our current iEval system (**see Appendices A-E**). We debated which questions were most helpful for providing formative feedback to instructors and evaluating teaching effectiveness, and weighed the balance of evaluating student satisfaction with the professor versus evaluating the learning experience (e.g. gains in knowledge, course content, and instruction).

Subcommittee A spent considerable time discussing the problem of bias against women faculty and faculty of color in student evaluations of teaching, both in the instrument itself and in the way that evaluations are deployed in merit and promotion.¹ We identified the problem of bias as among the most serious flaws in our current SET and its uses on our campus. UCR has the lowest percentage of women faculty in the UC system (at around 38%) and the percentage of women faculty at the rank of Full Professor is about 25% below the national average (the national average is only 25% and UCR is 25% below this). We discussed various mechanisms for minimizing bias.

With this in mind, we support the **UCTLCR** conclusions on bias (reproduced in modified version here)²: The research suggests that bias is pervasive and that women faculty and faculty of color are significantly disadvantaged in academic personnel reviews by these surveys, particularly when institutions rely heavily on quantitative

¹ <https://www.insidehighered.com/news/2016/01/11/new-analysis-offers-more-evidence-against-student-evaluations-teaching>
<https://www.scienceopen.com/document/vid/818d8ec0-5908-47d8-86b4-5dc38f04b23e>

² <https://senate.universityofcalifornia.edu/files/reports/kkb-divs-teaching-evaluation-task-force-report.pdf>

measures to assess faculty performance. Numerous studies going back 40 years and continuing today have established widespread bias in SETs (for recent reviews of the literature, see Basow et al., 2013; Spooren et al., 2015). Most scholarly attention within this literature has been paid to gender bias, and specifically how gender influences students' expectations for, and criticisms of, teaching performance. One leading study showed that when two instructors in an online course disguised their gender, with each instructor operating under two different gender identities, students rated the "male" identity significantly higher than the "female" identity, regardless of the instructor's actual gender (MacNeill et al., 2015). Such role expectations have special relevance in particular classroom contexts. Female instructors face less bias in small classes where individual interaction with students is the norm but are at a disadvantage in larger, more impersonal classes. Bias based on race and ethnicity has also been identified in the literature (Anderson and Smith, 2005; Bavishi et al., 2010; Smith and Hawkins, 2011).

We suggest the following immediate reforms:

- remove comparative statistics for Departments and Campus
- include a bias statement to be included at the beginning of evaluations

Overall Recommendations

Continue use of Student Evaluation of Learning

Our subcommittee agrees that student feedback is essential to teaching and learning on our campus and that there must be a mechanism for students to communicate their experience. This is particularly true at UCR where many of our students are first-generation, minority, and/or from socioeconomically disadvantaged backgrounds. These factors may shape the student learning experience in ways that are unfamiliar to many faculty. SETs can provide a bridge to communicate those challenges and experiences. We believe our students have an essential role to play not as consumers but as active learners who are engaged in their education. Their reported experience in the classroom is valuable information that can help support UCR's mission by influencing instructional approaches.

We therefore recommend the continued use of a substantively revised Student Evaluation of Learning survey instrument, while acknowledging considerable faculty frustrations with our SET to date. We believe that the SET can be improved to become more fair and useful, but acknowledge it will always be an imperfect instrument. Therefore, to reduce our campus' current overreliance on SETs for the purposes of merit and promotion, we also recommend that additional forms of evidence of teaching effectiveness be widely adopted.

Maintain numerical scales

Our subcommittee agrees that given the size of UCR's student body that our revised survey instrument must continue to include some scale-based questions, but we support the recommendations of the UC Centers for Teaching and Learning that these be changed to frequency scales (often/sometimes/rarely) or scales of agreement

(UCTLCR p6). We also recommend that our campus maintain the five-point scale, as evidence continues to suggest this is the most effective.³

However, in keeping with the **UCTLCR 2020 Report** we advise against using this data to create campus and department medians against which individual faculty are compared. UCTLCR recommends that campuses should not compare one instructor's scores to departmental averages, for example (p4). We refer to the guidance of Subcommittee B on this matter.

Substantially revise the current iEval survey

Our subcommittee recommends the substantive revision of our iEval questions both to make them a better measure of student learning outcomes and a more useful instrument to faculty as they strive to improve their teaching. Rather than simply removing problematic questions, each question adopted in our new system should reflect the best and most recent research. We endorse the guidelines in developing SET that appear on pp 2-3 of the UCTLCR Report.

Below we make nine specific recommendations for an iEval revision. Among these suggestions we provide several examples to illustrate what a modified iEval question might look like. However, we do not feel that our Ad Hoc committee has the research expertise to determine which questions are most effective measures and least prone to bias. We recommend the campus convene a group of faculty whose research is in the field of teaching, learning, education, and pedagogy to select, create, and/or identify the most effective questions drawing on the most current research.

1. Our campus should fully prepare and equip our students to participate in the evaluation of teaching as described in **UCTLCR 2020 Report** (p6). This preparation should include providing information and resources to students to learn about the importance of student feedback, providing examples of helpful and unhelpful feedback, and creating multiple opportunities for students to learn about the importance of evaluations. We add here that students need to learn about the problem of bias in SET.
2. There should be fewer and more precisely crafted questions. This may reduce evaluation fatigue, leading to higher participation rates and more meaningful responses from students. If a question does not **often** provide useful information to faculty or P&T evaluation committees, it should be removed to promote more careful consideration of questions with higher assessment value.
3. Those questions most prone to bias should be excluded. The current iEval questions #10, #11, #12, #13, and possibly others, fall into this category. Question #13 regarding the "overall ranking of the instructor" is especially of concern since many departments and P&T evaluation bodies rely on it centrally.

³ Revilla, Melanie A., et al. "Choosing the Number of Categories in Agree–Disagree Scales." *Sociological Methods & Research*, vol. 43, no. 1, 2013, pp. 73–97., doi:10.1177/0049124113509605.

4. iEval questions should focus on student learning rather than instructor performance. **UCTLCR 2020 Report** provides examples of student experience centered questions that can help assess the effectiveness of a class. We note that a “Student Evaluation of Teaching” implies a one-way relationship in which the instructor delivers a product and the student evaluates their satisfaction as a consumer. We propose a transition to a “Student Evaluation of **Learning**” that aims to rebalance this relationship by acknowledging that learning is a two-party process. Some committee members felt that iEval questions #6-#13 fall into the category of instructor performance and should be excluded. Other committee members felt strongly that it was important to preserve some evaluation of instructor performance.
5. Consider including the instructor’s stated learning outcomes or learning objectives for the course as a component of the student evaluation. For example:

For each of the following course learning objectives (if provided), please evaluate the gains that you have made this quarter.

Minimal=1 Moderate=3 Strong=5

Course Objective 1: <list instructor-provided course objective here> 1 2 3 4 5

Course Objective 2: <list instructor-provided course objective here> 1 2 3 4 5

Course Objective 3: <list instructor-provided course objective here> 1 2 3 4 5

Course Objective 4: <list instructor-provided course objective here> 1 2 3 4 5

6. Narrative comment fields should be retained, but the iEval survey should provide carefully crafted prompts that guide the scope of student comments rather than simply having an open-ended Comments field. This shifts the tone of the evaluation from a consumer satisfaction survey to a more formative evaluation, and may provide useful information for faculty to understand how students engage with the course. For example:

What approaches or materials did you find most effective or successful in aiding your learning in this course? Please describe approaches that you used as well as those of the instructor.

<short answer box>

What changes could you have made and/or the instructor have made to improve your ability to achieve the learning goals of this course?

<short answer box>

7. A statement or “preamble” about bias should be presented to students at the beginning of the survey (see example below). Evidence suggests that such statements are partially effective in mitigating bias. This change could be implemented easily and immediately as a stopgap measure. We recommend that additional strategies are developed to educate both students and faculty about the prevalence of bias in SETs.

“Student evaluations play an important role in ongoing instructional improvement and institutional reviews of faculty teaching effectiveness. The University of California, Riverside recognizes that student evaluations of teaching may be influenced by unconscious and unintentional biases about the gender and race of the instructor. Women faculty and faculty of color are systematically rated lower in their teaching evaluations than white male faculty, even when there are no actual differences in instruction or what students have learned.

As you fill out the course evaluation, please focus your opinions on the content of the course, the course materials, and the quality of your learning.”

8. Consider using the Student Evaluation of Learning to gather information about other obstacles to learning our students face that are beyond the control of the instructor. This information would be shared with administration to guide campus development and resource allocation. It may also be used by P&T evaluation committees to identify factors that may unduly influence student perceptions of faculty teaching effectiveness in a course. For example:

Was the university-provided infrastructure (for example, classroom facilities and equipment, external student resources, class size, number of teaching assistants) appropriate to promote positive learning outcomes in this course?

<short answer box>

9. The placement of prompts for written comments relative to scale-based questions may influence numerical ratings. It is possible that providing written comments first, particularly if the student is invited to reflect on their role in the learning process, may provoke a more thoughtful and less biased consideration of scores on scale-based questions. Alternatively, specific criteria that are raised by scale-based questions may provide a useful framework for composing written comments. We recommend that a campus study be performed to evaluate this hypothesis.

Members of the Subcommittee A Task Force

John Heraty, Subcommittee A Chair, Professor, Entomology
Michelle Bloom, Professor, Comparative Literature and Languages
Kevin Costello, Associate Professor of Teaching, Math
Jennifer Scheper Hughes, Associate Professor, History
Suveen N. Mathaudhu, Associate Professor, Physics
Dave Nelson, Committee Chair, Associate Professor, Botany and Plant Sciences
David Gracey, Associate Vice Chancellor, Information and Technology Services
Stefano Lonardi, Professor, Computer Science & Engineering (2018-19)
Traise Yamamoto, Associate Professor, English (2018-19)
Teri Eckman, Interim Executive Director, Student Information Systems
Israel Fletes, Director of Academic Technology (2018-19)

REFERENCES

- Adams, C.L. (2012). Online measures of student evaluations of instruction. In M. E. Kite (Ed.), *Effective evaluation of teaching: A guide for faculty and administrators*. Retrieved from the Society for the Teaching of Psychology web site: <http://teachpsych.org/ebooks/evals2012/index.php>
- Anderson, K. J., & Smith, G. (2005). Students' preconceptions of professors: Benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences*, 27(2), 184-201.
- Ballantyne, C. (2003). Online evaluations of teaching: An examination of current practice and considerations for the future. *New directions for teaching and learning*, 2003(96), 103-112.
- Basow, S., Codos, S., & Martin, J. (2013). The effects of professors' race and gender on student evaluations and performance. *College Student Journal*, 47(2), 352-363.
- Bavishi, A., Madera, J. M., & Hebl, M. R. (2010). The effect of professor ethnicity and gender on student evaluations: Judged before met. *Journal of Diversity in Higher Education*, 3(4), 245.
- Benton, S. L., & Ryalls, K. R. (2016). Challenging Misconceptions about Student Ratings of Instruction. IDEA Paper# 58. *IDEA Center, Inc.*
- Becker, W. E., Bosshardt, W., & Watts, M. (2012). How departments of economics evaluate teaching. *The Journal of Economic Education*, 43(3), 325-333.
- Berk, R. A. (2012). Top 20 strategies to increase the online response rates of student rating scales. *International Journal of Technology in Teaching and Learning*, 8(2), 98-107.
- Bernstein, D. J., Jonson, J., & Smith, K. (2000). An examination of the implementation of peer review of teaching. *New Directions for Teaching and Learning*, 83, 73-86.
- Braskamp, L. A., Ory, J. C., & Pieper, D. M. (1981). Student written comments: Dimensions of instructional quality. *Journal of Educational Psychology*, 73(1), 65-70. doi.org/10.1037/0022-0663.73.1.65
- Brookfield, S. (1998). Critically reflective practice. *J. Contin. Educ. Health Prof.*, 18: 197-205. [doi:10.1002/chp.1340180402](https://doi.org/10.1002/chp.1340180402)
- Boring, A., Ottoboni, K., & Stark, P. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*.
- Chism, N. V. N., & Chism, G. W. (2007). *Peer review of teaching: A sourcebook*. Bolton, MA: Anker Pub. Co.
- Crews, T. B. & Dylan F. Curtis, D. F. (2011). Online course evaluations: Faculty perspective and strategies for improved response rates. *Assessment & Evaluation in Higher Education*, 36:7, 865-878, DOI: 10.1080/02602938.2010.493970
- Dommeyer, C. J., Baum, P., Hanna, R. W., & Chapman, K. S. (2004). Gathering faculty teaching evaluations by in-class and online surveys: their effects on response rates and evaluations. *Assessment & Evaluation in Higher Education*, 29(5), 611-623.
- Gillman, R. A., Hensel, N. H., Salomon, D. A., & Wilhite, S. C. (2018). *REFINING THE PARADIGM*.
- Hoel, A., & Dahl, T.I. (2018). Why bother? Student motivation to participate in student evaluations of teaching. *Assessment and Evaluation in Higher Education*, DOI: 10.1080/02602938.2018.1511969.
- Hornstein, H. A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education*, 4(1), 1304016.

- MacNeill, L., Driscoll, A., & Hunt, A. N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Innovative Higher Education*, 40(4), 291-303.
- Nutly, D. D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment and Evaluation in Higher Education*, 33(3, June), 301–314.
- Peterson, D. A., Biederman, L. A., Andersen, D., Ditonto, T. M., & Roe, K. (2019). Mitigating gender bias in student evaluations of teaching. *PloS one*, 14(5), e0216241.
- Seldin, P. (2006). *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. Bolton, MA: Anker Pub. Co.
- Smith, B. P., & Hawkins, B. (2011). Examining student evaluations of Black college faculty: Does race matter?. *Journal of Negro Education*, 80(2).
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, 83(4), 598-642.
- Stark, P. B., & Freishtat, R. (2014). *An evaluation of course evaluations*. doi:10.14293/S2199-1006.1.SQR-EDU.AOFRQA.v1Stark. Retrieved from Science Open:<https://www.scienceopen.com/document/vid/42e6aae5-246b-4900-8015-dc99b467b6e4?0>
- Svinici, M. D. (2001). Encouraging your students to give feedback. *New Directions for Teaching and Learning*, 87, 17-24.
- "Statement on Student Evaluations of Teaching," American Sociological Association, August 2019
- Uttl, B., White, C. A., & Gonzalez, D. W. (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.
- Worthington, A.C. (2002). The impact of student perceptions and characteristics on teaching evaluations: A case study in finance education, *Assessment and Evaluation in Higher Education*, 27(1), 49–64.

APPENDIX A - Most Relevant Recommendations from the 2019 UC Teaching and Learning Centers Report

- R1** Campuses should rely on multiple methods to evaluate teaching effectiveness, balancing student evaluation of teaching (SET) scores with other measures (i.e. peer observation, self-assessment).
- self-assessment should provide a reflection on their teaching development (focusing on growth rather than attainment).
 - peer observations should be separated into assessment (internal) versus evaluation (external merit review). Note that peer observation is also subject to the same sort of unconscious bias that troubles SET.
 - Encourage use of teaching portfolios.
- R2** Campuses should adopt the broadest language possible to more fully capture the kinds of bias that may be systemic in teaching evaluations.
- Eliminate heavily gendered criteria such as “helpfulness” and “warmth” from survey questions.
 - Change name to Student Experience of Teaching Surveys
 - Bias training and education about/on the importance of the evals are important at all levels.
 - Institute a process for faculty who wish to respond to discriminatory comments.
- R3** Campuses should adopt a more developmental approach to teaching (i.e., create opportunities for faculty to experiment and innovate in their courses without fear of “failing,” rewards teaching professional development.
- R4** Campuses need to be willing to spend more time evaluating teaching effectiveness and expect instructors to spend more of their (finite) time improving their teaching.
- R5** Campuses should clarify what aspects of teaching SET the questions are attempting to evaluate.
- Rewrite SET questions to reflect student experiences in the classroom (“the instructor defined important concepts”)
 - Shift from numerical scales to frequency scales.
- R6** Specific questions that prompt students to consider different aspects of a course and its instruction should precede more global questions.
- R7** SETs should include open-ended questions (comments) that ask students to describe their experience in the course, what aspects facilitated their learning, and how it could be improved
- R8** If quantitative measures (such as Likert scales) are used for response items, they should be associated with questions about students’ experience of learning rather than asking questions outside of their experience. For instance: “How much would you say you have learned in this course?” is a question that focuses on students’ own experiences.
- R9** Provide professional development opportunities for faculty to improve their teaching.

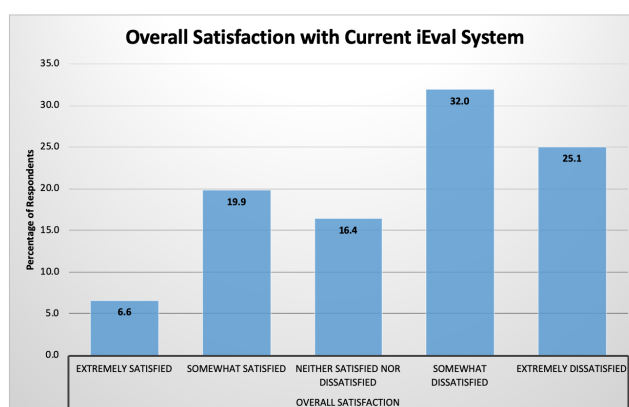
APPENDIX B - Summary of 2020 UCR campus survey on iEval

The iEval Faculty Opinion Survey was conducted by the Ad Hoc Committee on Evaluation of Teaching with the assistance of Dr. Omar Safie, Director of Evaluation & Assessment (*ex officio* member). The survey was distributed to all instructors at UCR, and was left open from May 18, 2020 to June 18, 2020, with two reminders sent to faculty who had not yet completed the survey. Respondents completed surveys anonymously and their responses will be kept confidential within the committee. A total of 351 respondents completed all or part of the survey.

The majority of faculty surveyed were from CHASS, followed by CNAS. CHASS and CNAS have the largest populations of majors and faculty at UCR. Combined, faculty from CHASS and CNAS comprise nearly 80% of survey participants. BCOE, GSE, SOB accounted for 16.3% of the responses. The faculty completing the survey were nearly evenly split between Man and Woman cisgender with a significant percentage of respondents declining to state.

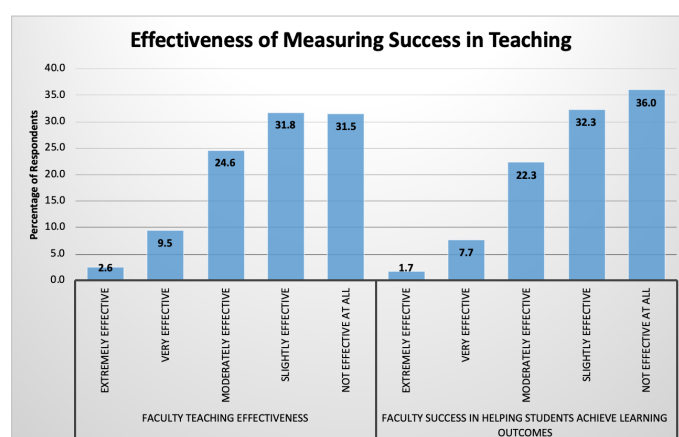
Based on IPEDS definitions of ethnicity and race, the majority of participants were not of Hispanic/Latino ethnicity and were majority White. However, it is important to note that a quarter of faculty completing this survey declined to state their ethnicity or race.

The majority of faculty indicated dissatisfaction with the current iEval system.



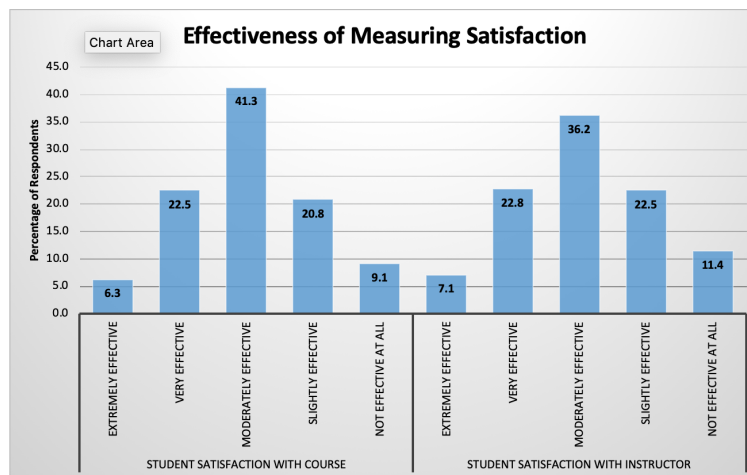
Effectiveness of Measuring Success in Teaching

In addition to their overall dissatisfaction with the current iEval system, a majority of faculty indicated that the current iEval system **was not very effective at measuring teaching effectiveness or faculty success** in helping students achieve learning outcomes. This indicates a need to refocus iEval on factors that measure teaching effectiveness and student success, such as student feedback on teaching methodologies and a focus on achievement of outcomes.



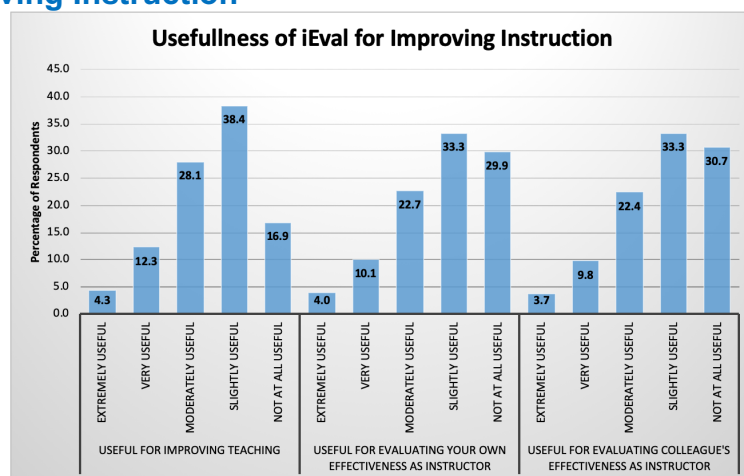
Effectiveness of Measuring Satisfaction

The largest percentage of faculty surveyed landed right in the middle, indicating the current iEval system was moderately effective at measuring satisfaction with the course or with the instructor



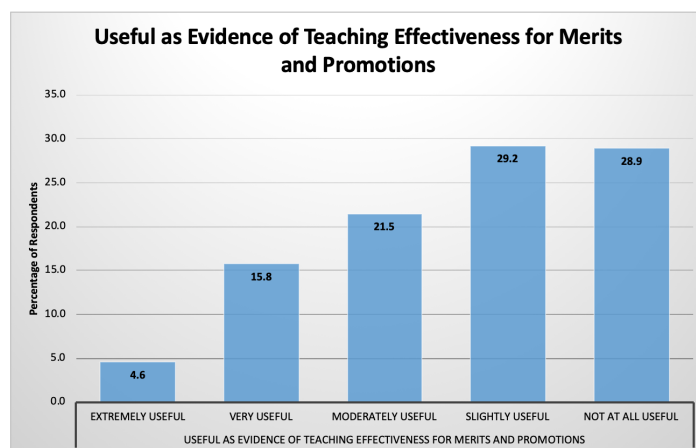
Usefulness of iEval for Improving Instruction

A majority of faculty did not find the current iEval system useful, or only slightly useful, in evaluating their own or a colleague's effectiveness as an instructor. While more faculty indicated that the current iEval system is moderately to extremely useful for improving teaching, a general review of comments indicated that this is dependent on the open-ended comments from students



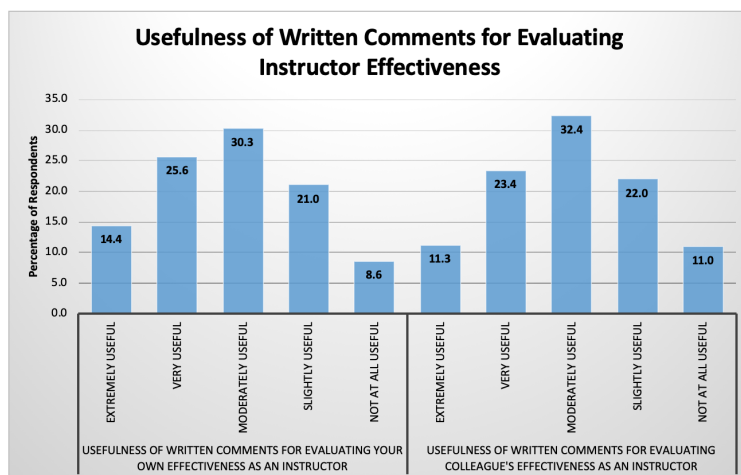
Usefulness as Evidence of Teaching Effectiveness for Merits and Promotions

The majority of faculty surveyed indicated that the current iEval system was only moderately useful, at best, as evidence of teaching effectiveness for merits and promotions.



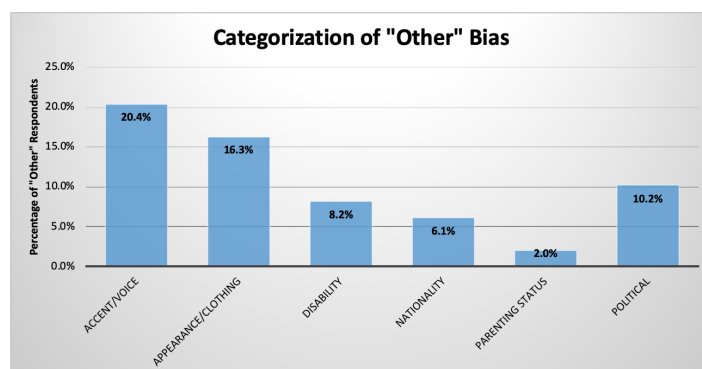
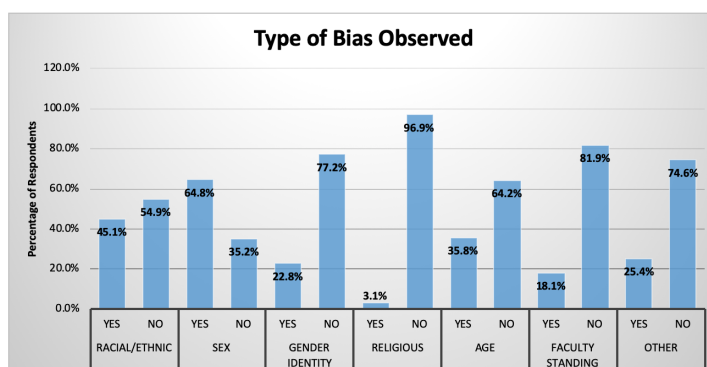
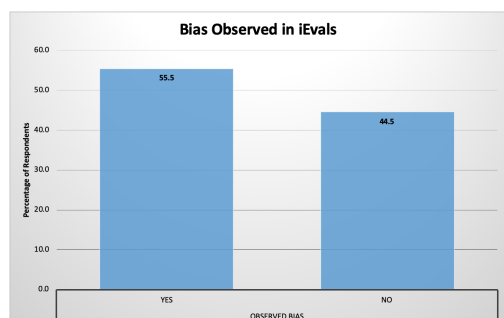
Usefulness of Written Comments for Evaluating Instructor Effectiveness

the majority of faculty indicated that they were moderately to extremely useful for evaluating their own effectiveness as an instructor. However, a nearly evenly split percentage of faculty indicated that the written comments were useful for evaluating a colleague's effectiveness as an instructor



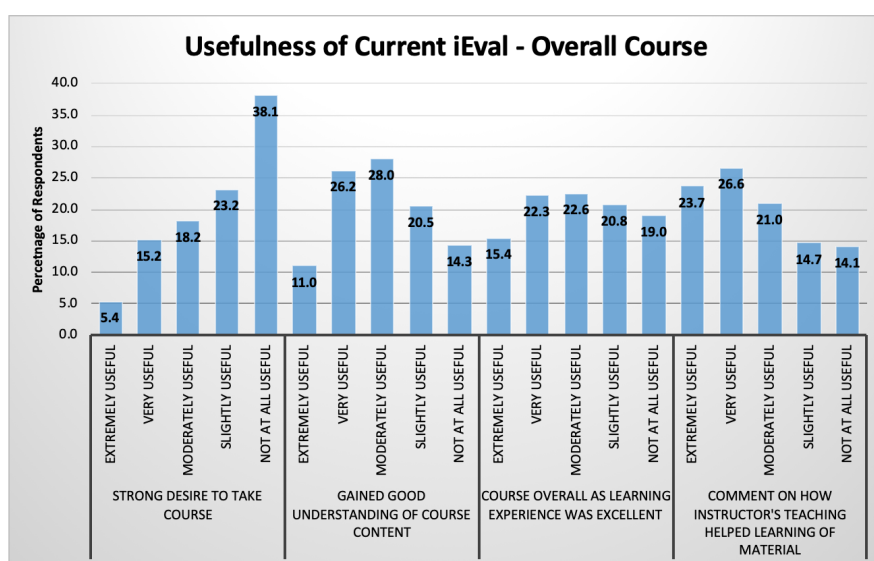
Bias in Current iEval

The majority of faculty indicated that they observed bias in the current iEval system.



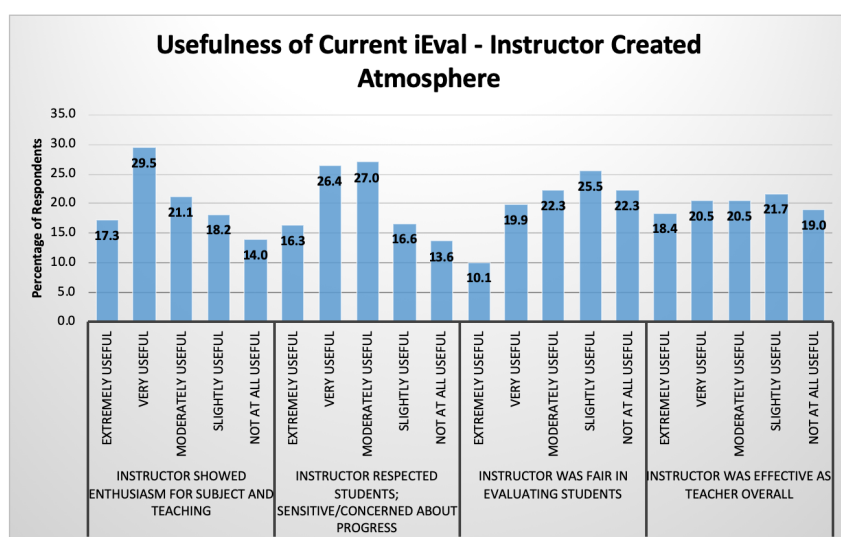
Usefulness of Current iEval – Overall Course Questions

The majority of faculty indicated that asking students if they had a strong desire to take a course is not at all useful or only slightly useful. Given that some courses are required for majors, such a question does not seem to be useful for faculty seeking information to improve their practice. In addition, asking students if a course was an excellent learning experience also may not be useful to faculty. However, more faculty found useful the questions focused on whether the student gained a good understanding of course content and the comments on how an instructor's teaching helped the learning of material. This lends support to a need to shift from satisfaction, general questions, and comparative rankings, to more focused questions on learning and outcome achievement that provide useful information for faculty to improve their instruction.



Instructor Created Atmosphere

Faculty surveyed indicated that faculty demonstrating enthusiasm for the subject and teaching was useful with nearly 30% indicating very useful. In addition, faculty surveyed also indicated that instructor respecting and being sensitive to and concerned about student progress were also important. However, nearly 50% of faculty surveyed indicated that the question related to instructor being fair in evaluating students was only slightly or not at all useful.



Overall Suggestions Based on Survey Results

The suggestions below are preliminary and are based on a combination of findings from this survey and the UC Centers of Teaching and Learning November 2019 report on Recommendations for Evaluating Teaching Effectiveness. These suggestions are meant to move the conversation forward. They are not conclusive nor are they prescriptive.

1. iEval Content

- a. Exclude questions that focus on satisfaction.
- b. Exclude questions that focus on fairness.
- c. Include questions that focus on instructor implementation.
- d. Include questions that focus on usefulness/applicability of course materials, but excludes exams.
- e. Include questions that focus on achievement of course outcomes.
- f. Include questions that focus on faculty created atmosphere.
- g. Include open-ended questions that are focused on improving instruction and not left completely open.

2. iEval Use

- a. Consider how iEval findings are used. Specifically, how should iEval findings be used as a way to evaluate teaching effectiveness and for merits and promotions?
 - i. One solution is to ensure that iEval Content is focused on teaching effectiveness and not satisfaction.
 - ii. Another is to allow faculty the opportunity to reflect on and speak to how they used iEval reports to improve their practice as instructors. This allows for faculty to share not only the context within which the course was implemented and how they managed to “close the loop.”
- b. Consider removing comparative rankings in the iEval report. Given the vast differences between courses across the campus and even within the same discipline, the usefulness of this aspect of the report is limited.

3. iEval Implementation

Consider providing a core set of questions for all faculty, as well as a set of questions that individual faculty or departments can “add” for their iEvals. This would allow for flexibility for faculty and programs to not get a core set of results from iEval, as well as results that may be more relevant/useful to specific cases.

Appendix C - List of current iEval questions for context

1. I had a strong desire to take this course
2. I attended class regularly
3. I put considerable effort into this course
4. I gained a good understanding of the course content
5. I normally spent at least two hours preparing for each hour of class
6. Instructor was prepared and organized
7. Instructor used class time effectively
8. Instructor was clear and understandable
9. Instructor exhibited enthusiasm for subject and teaching
10. Instructor respected students; sensitive to and concerned with their progress
11. Instructor was available and helpful
12. Instructor was fair in evaluating students
13. Instructor was effective as a teacher overall
14. The syllabus clearly explained the structure of the courses
15. The examinations reflected the materials covered during the course
16. The required readings contributed to my learning
17. The assignments contributed to my learning
18. Supplementary materials were informative (e.g. films, slides, videos, demonstrations, guest lectures, iLearn, web pages, etc)
19. The course overall as a learning experience was excellent

Comment:

Appendix D - Some potential questions suggested from the faculty survey

- Describe aspect(s) of the course that you found beneficial to your learning.
- Describe one thing YOU (the student) can do to improve your learning experience in the course. Describe one thing I (the instructor) can do to improve your learning experience in the course.
- Did this course lead you to want to take more courses in this field of study?
- Have you grown in your knowledge of the course subject matter?
- How would you rate the overall course experience?
- How would you rate your own contribution to the course? Do you come on time, pay attention, stay off your phone, actively contribute to you own learning? Do you treat the learning environment with respect and professionalism?
- The instructor appeared to reflect and tried to adapt their teaching when things didn't go as planned or when students were struggling.
- What's Been most Helpful to Your Learning? Please select the teaching element that has been most helpful to your learning, and then provide a detailed written comment about what worked well and why. 1. The inclusiveness of this course. 2. Support from the instructor. 3. Feedback from the instructor. 4. The level of challenge of this course. 5. The quality of course materials. 6. The clarity of instructions for assignments and grading. 7. The use of active learning practices. 8. Interactions between students in this course. 9. Instructor communication in this course. 10. The organization of this course. 11. The relevance of the course content. 11. The assignments or projects in this course. 12. The accessibility of this course. 13. None of the elements above are helpful to my learning.
- Do you feel the course was a valuable experience in helping to develop your own thinking/direction/approach on the subject?
- I have considered the possibility that personal biases around race, gender identity, age and sexual orientation may impact my responses to this survey.
- The instructor explained concepts clearly.

Appendix E - Examples of negative comments provided in the faculty survey

- Abandon evaluations as other institutions have. They do not work.
- Allow ONLY positive student comments to be included in merit letters; no quoting of negative comments.
- Are you aware that faculty who are women and/or persons of color are rated lower across the board in faculty evaluations, nationwide?
- Get rid of evaluations. They are largely useless. And, since they are optional, they only take a small sampling from the class, which gives very unrealistic data.
- I do not believe our teaching should be numerically evaluated — every study indicates that these formats are particularly vulnerable to bias.
- I wouldn't recommend anything since the SET approach is not valid and has come under extreme review with all kinds of empirical support that the internal validity of the instrument is poor at best and deeply misleading and misused at worst. They suffer from significant gender discrimination.
- If it is not mandatory that all students complete ieval then they are of no use. It does not give an accurate portrait.
- Student evaluations should never be used alone to evaluate a professor. I would highly recommend that an unbiased committee of peers also be involved in evaluations. And why must the University evaluate a professor every single quarter. It leads to grade inflation as studies have shown us for a very long time. Everyone knows who the good professors are and to evaluate each quarter is folly. If there is a problem evaluations are necessary but otherwise why? We all know professors who get great evaluations and also noticed they give mostly A's and few C's and little else. This is not always the case of course but it is a dirty little secret that to get better evaluations the trick is to lower your standards. We all know that and the University knows that as well as this topic has been debated for many years. If you really wish to know about the teaching abilities of a professor look at their syllabus and the quality of their exams. Come to their classes and listen to their lectures. This would benefit both the professors and especially the students. But I do not expect any changes as I said above that it has long been known and researched again and again that teaching evaluations by students alone have unintended consequences. My question is: why are you bothering with a survey? Let me end by saying that some of my best teachers during my school years were not appreciated by me until years later. They made me work hard, they challenged my convictions and the like. And I feel so sad that I didn't appreciate them in my youth. In fact, they are responsible in large part for my success as they forced (a rather lazy student) to work hard and earn my grades and to love the world of ideas.
- Wrong approach! The problem is the inherent biases of ieval, not the content of individual questions.

SUBCOMMITTEE B (best practices to evaluate teaching)

Charge

Provide recommendations to the campus on how to interpret and use teaching evaluations responsibly for merit/promotions. Address the following questions:

- *What biases (or other factors) should department chairs or ad hoc merit/promotions committees keep in mind when using teaching evaluations?*
- *How should the data be presented so they are more intuitive to interpret, without encouraging problematic uses or abuse?*
- *How to incorporate for class size, level (lower/upper division, grad/undergrad, required/elective), type (lecture, on-line, hybrid, flipped), etc. in the interpretation of teaching evaluations?*
- *How to implement the instructions (APM 210-1-d) to Review Committees that Advise on Actions Concerning Appointees in the Professor and Corresponding Series when considering the teaching record of the candidate's performance.*
- *How to incorporate more comprehensive models of teaching evaluations (e.g., teaching portfolio)?*

Item 1: *What biases (or other factors) should department chairs or ad hoc merit/promotions committees keep in mind when using teaching evaluations?*

Overall, transparency and fairness should guide all forms of personnel evaluation. Furthermore, evaluators should reflect upon their personal assumptions and biases in their evaluative role. In light of this, we have three primary areas for recommendations:

Academic Personnel Office (VPAP & APO) and Committee on Academic Personnel (CAP) Direction

At the central administrative and Academic Senate level, evaluators need to communicate their role in academic personnel evaluation to Colleges and Departments, so that all academic personnel who undergo review are clear about these roles. Part of this communication, consistent with transparency and fairness, should indicate (a) the separate and distinct roles of the VPAP and CAP; (b) the goals and possible biases of these entities (e.g. 1, that the VPAP reports to the Provost and Chancellor and has a legal responsibility to the university, and thus has some common level of interest with the university's administration; e.g. 2, that CAP members ascribe to some shared sense of academic merit and productivity that may retain some measure of disciplinary biases); and (c) the emphases each party places upon a faculty member's overall performance (e.g., research productivity, resource acquisition, teaching and mentorship, and the quantity and quality of university and professional service).

Policy Guidance for Department Chairs, Associate/Divisional Deans, and Deans

At this level, there is a need for guidance on evaluation practices that constitute both intentional and unintentional bias. This guidance could include definitions of bias, forms of biases, and familiar biases evidenced across teaching evaluations (including student evaluations of teaching). For example, forms of bias in evaluation include weighted (in)appropriate disciplinary expertise, identity category bias/social stereotypes, teaching course assignment biases, and quantitative/qualitative research biases.

Interpretative directions/guidance for individual Faculty members and Department Chairs in deliberation meetings

For department chairs, there is a need for guidance with respect to the evaluation of teaching, particularly with reference to the Call and APM, so that matters outside of evaluation policy do not enter into deliberations or review decisions. For example, there is great inconsistency across departments about whether it is permissible to discuss family leave in personnel deliberations; this should be clarified for both candidates and department faculty members long before such meetings. Furthermore, there is a need for department chairs, and those chairing departmental review committees, to inform departmental members about what constitutes conscious and implicit bias in teaching and mentoring. Department chairs and departmental members need to become cognizant of relevant sections of the APM that can be loaded with bias: for example, "through *objective...* appraisal" (APM 210-1a) and "ability to organize material and to present it with *force and logic*" (APM 210-d1).

Item 2: *How should the data be presented so they are more intuitive to interpret, without encouraging problematic uses or abuse?*

Currently, UCR's numeric SET reports consist of 19 items, each containing 17 columns of data: the distribution of student responses, followed by the mean, median, and standard deviation at a course, department, and campus level, as well as department and campus-wide percentile rankings. In practice, many instructors and evaluation committees ignore most of those numbers and simply review the columns that are easiest to take in at a glance: the mean and the percentiles. Unfortunately, these are not good representations of student responses. At the same time, the volume of data on UCR's SET reports, and the increased precision to which the data are being calculated, invite a degree of comparative analysis that may not be warranted.

Some changes to how the data are presented in SET reports—and what data are presented—can guide reviewers to the most useful information, reign in the temptation to overanalyze, and avoid what Philip B. Stark calls “quantifauxcation.”

Item Means

Prior to Fall 2016, the item means on SET reports were rounded to one decimal place. Since then, the means have been rounded to two decimal places. However, student responses do not support this level of precision. A mean of 4.3 is little different from a 4.4, and variations in the mean up to ± 0.3 or ± 0.4 might not be significant ([Cashin rec 23](#), [Linse 6.3](#)). If item-by-item means are to be presented in SET reports, they should be rounded no further than one decimal place (Cashin rec 22).

Percentiles

Few campuses include percentile rankings in SET reports. For instance, among eight UC campuses surveyed so far, only UCR includes percentile rankings, and UCR has had difficulty making them work well.

Prior to Fall 2016, percentile ranks were assigned based on an item's mean rounded to one decimal place. Since then, they have been assigned based on an unrounded mean. This was done to reduce “clumps” in the percentile rankings and increased the number of ranks from roughly 25 to roughly 500. This is problematic, given the potential variation in means discussed above, but there are other considerations that further discourage the use of percentiles:

- Because student responses tend to be highly positive—typically in the 4-5 range—the percentile ranks re-cast a narrow range of mean values onto a 100 point scale. This tends to magnify insignificant differences. For instance, a 0.1 shift in an item's mean may translate to a jump of 10 or more percentile points. While the former is typically perceived as inconsequential, the latter can have the appearance of a significant “real” change in overall student response.
- The percentiles are frequently misinterpreted. As the 7 Sep 2016 memo “Changes to iEval Effective Fall 2016” notes, an item can have largely positive ratings (e.g. 85% positive) and still have a low percentile rank (35th percentile).

Yet in practice, when the percentile rank is low, reviewers have frequently interpreted the positive ratings as a sign of “poor performance.”

- Given the potential variation in the mean value of student responses, confusion about what SETs measure, and concerns about the degree to which gender, ethnicity, age, course subject, class size, teaching innovation, etc. can affect student responses, it is counterproductive to rank SET results in a department or across campus.

In short, the percentile rankings, whether intended or not, invite reviewers to do a kind of analysis that they should not be doing via SETs ([Linse 6.6](#)). The percentile rankings should be discontinued.

Department and Campus Statistics

SET reports at a number of universities, such as UCI, UCSC, and UCSD, show responses only for the course itself. Those at other universities, such as UCB and UCSB, add departmental statistics, usually in the form of item means. Fewer, such as at UCSB, add campus-wide statistics. This lack of consensus about what to report is reflected in the literature. Some, such as Cashin, argue that one cannot meaningfully interpret the responses without comparison. Others, such as Freishtat, Linse, and Stark argue that SETs are so subjective and confound so many variables that it makes little sense to compare them.

The disagreement may arise from the fact that SET reports have two audiences: instructors and evaluation committees. An instructor might wish to see the department and campus means as a rough way to contextualize responses for a course, but an evaluation committee might not, given the [low correlation](#) between ratings and teaching effectiveness. In the committee this was summed up in one question: How relevant are these numbers? This merits further discussion.

Open-Ended Comments

There have been concerns about open-ended comments, particularly with regard to their variability, relevance, and appropriateness. While they can often yield interesting information, the general consensus is to look for trends in responses—with the usual caveat that open-ended responses should also be taken with a grain of salt. One question in the committee was whether it would be worthwhile to run the comments through sentiment analysis software, as one might do in marketing research. The thought was that this might counterbalance the tendency of reviewers to latch onto individual comments, but there was also concern about how well this would work in practice. The collection of individual responses should be made available in any case.

Visualization

A Likert survey, such as UCR’s SET, is a collection of items that as a group provide information about an underlying phenomenon of interest; items should not be interpreted independently. It may be helpful, then, if SET reports explicitly grouped related items to encourage a holistic interpretation of student responses.

SET reports should also be presented in a way that foregrounds the distribution of responses ([Freishtat J.1 & J.2](#)). A table is the most frequent way to do this. Another option would be to [graph the responses](#). A grouped diverging bar chart can fit all of the items onto a single page, which may encourage reviewers to interpret responses holistically. It offers a quick, intuitive way to see the response distribution. And it visually emphasizes the overall positivity or negativity of responses. That may be all that one would reasonably wish to glean from numeric SETs. While it may be tempting to analyze the SETs in greater detail and complexity than this, one should instead look to other evidence of teaching effectiveness as recommended in APM 210-1-d.

A sample table and chart based on UCR's current SET are included below. Each gives a readable summary that foregrounds the distribution of responses without inviting over-analysis.

Summary of Recommendations

1. Round item means to one decimal place, if they are to be included.
2. Remove the department and campus percentiles from SET reports.
3. Consider whether or not the remaining department and campus statistics are relevant in SETs.
4. Consider using a grouped diverging bar chart as the primary method of presenting responses in SET reports, or alternatively, a table that foregrounds the distribution of responses.

Instructor Name
Faculty Evaluation Responses: 2019S, ANON 0000-0
(22 enrolled, 17 responses)

Student Profile

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)	Total	Mean	Median
1. I had a strong desire to take this course	1 (5.9%)	3 (17.6%)	5 (29.4%)	5 (29.4%)	3 (17.6%)	17	3.4	3.0
2. I attended class regularly	0 (0%)	0 (0%)	0 (0%)	5 (29.4%)	12 (70.6%)	17	4.7	5.0
3. I put considerable effort into this course	0 (0%)	0 (0%)	0 (0%)	9 (52.9%)	8 (47.1%)	17	4.5	4.0
4. I gained a good understanding of the course content	0 (0%)	0 (0%)	1 (5.9%)	9 (52.9%)	7 (41.2%)	17	4.4	4.0
5. I normally spent at least two hours preparing for each hour of class	0 (0%)	2 (11.8%)	3 (17.6%)	5 (29.4%)	7 (41.2%)	17	4.0	4.0

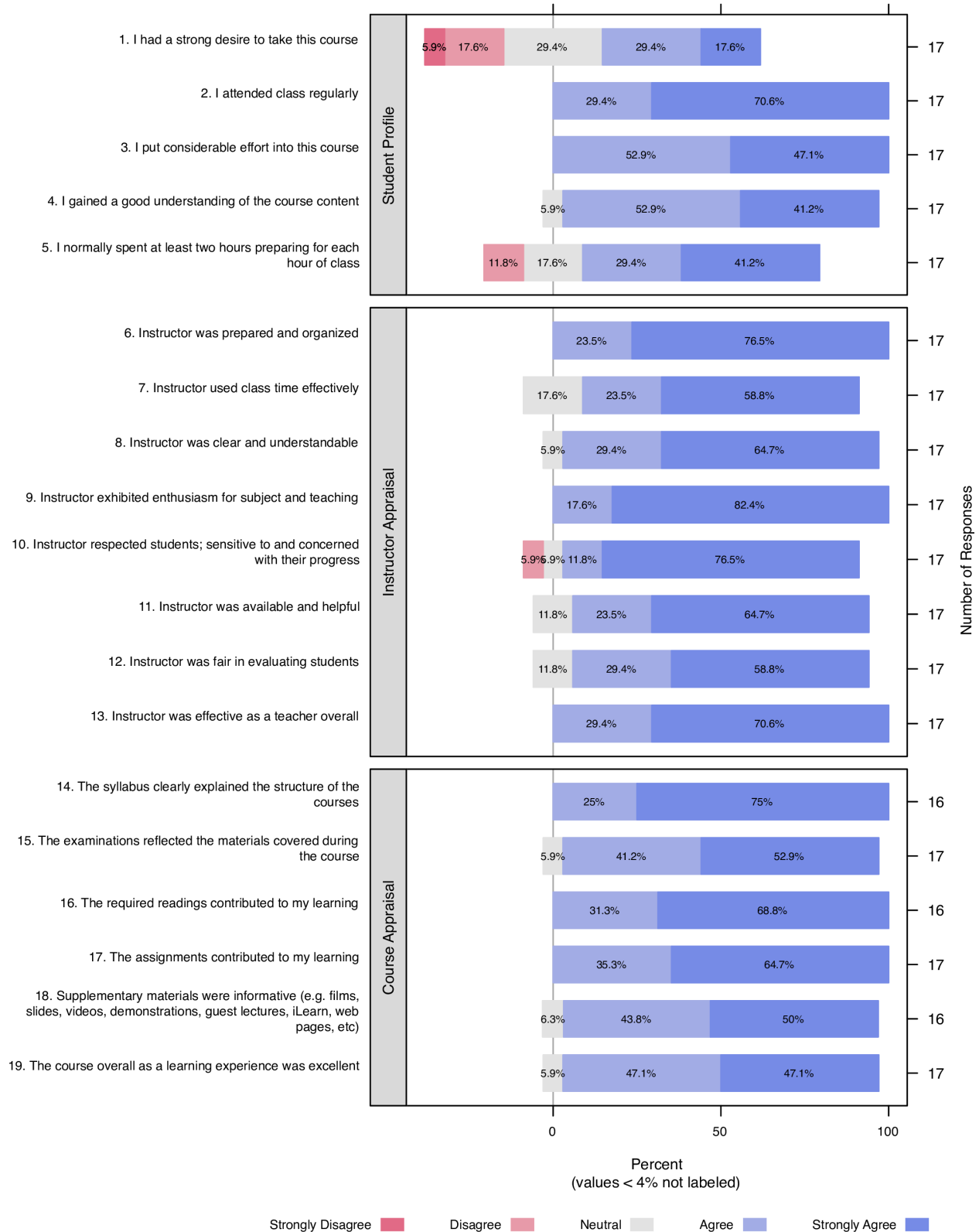
Instructor Appraisal

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)	Total	Mean	Median
6. Instructor was prepared and organized	0 (0%)	0 (0%)	0 (0%)	4 (23.5%)	13 (76.5%)	17	4.8	5.0
7. Instructor used class time effectively	0 (0%)	0 (0%)	3 (17.6%)	4 (23.5%)	10 (58.8%)	17	4.4	5.0
8. Instructor was clear and understandable	0 (0%)	0 (0%)	1 (5.9%)	5 (29.4%)	11 (64.7%)	17	4.6	5.0
9. Instructor exhibited enthusiasm for subject and teaching	0 (0%)	0 (0%)	0 (0%)	3 (17.6%)	14 (82.4%)	17	4.8	5.0
10. Instructor respected students; sensitive to and concerned with their progress	0 (0%)	1 (5.9%)	1 (5.9%)	2 (11.8%)	13 (76.5%)	17	4.6	5.0
11. Instructor was available and helpful	0 (0%)	0 (0%)	2 (11.8%)	4 (23.5%)	11 (64.7%)	17	4.5	5.0
12. Instructor was fair in evaluating students	0 (0%)	0 (0%)	2 (11.8%)	5 (29.4%)	10 (58.8%)	17	4.5	5.0
13. Instructor was effective as a teacher overall	0 (0%)	0 (0%)	0 (0%)	5 (29.4%)	12 (70.6%)	17	4.7	5.0

Course Appraisal

	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)	Total	Mean	Median
14. The syllabus clearly explained the structure of the courses	0 (0%)	0 (0%)	0 (0%)	4 (25%)	12 (75%)	16	4.8	5.0
15. The examinations reflected the materials covered during the course	0 (0%)	0 (0%)	1 (5.9%)	7 (41.2%)	9 (52.9%)	17	4.5	5.0
16. The required readings contributed to my learning	0 (0%)	0 (0%)	0 (0%)	5 (31.3%)	11 (68.8%)	16	4.7	5.0
17. The assignments contributed to my learning	0 (0%)	0 (0%)	0 (0%)	6 (35.3%)	11 (64.7%)	17	4.6	5.0
18. Supplementary materials were informative (e.g. films, slides, videos, demonstrations, guest lectures, iLearn, web pages, etc)	0 (0%)	0 (0%)	1 (6.3%)	7 (43.8%)	8 (50%)	16	4.4	4.5
19. The course overall as a learning experience was excellent	0 (0%)	0 (0%)	1 (5.9%)	8 (47.1%)	8 (47.1%)	17	4.4	4.0

Instructor Name
Faculty Evaluation Responses: 2019S, ANON 0000-0
(22 enrolled, 17 responses)



Item 3: *How to incorporate for class size, level (lower/upper division, grad/undergrad, required/elective), type (lecture, on-line, hybrid, flipped), etc. in the interpretation of teaching evaluations?*

The question of how to evaluate different categories of courses has been addressed at a number of different institutions. In our initial analysis, we have concluded that perhaps the best solution may be to differentiate mostly on class type, referenced as standard (i.e. lecture-based) and non-standard (utilizing technology and not requiring face-to-face interactions) formats. This has resulted in two recommendations for these two formats with the first recommendation giving faculty members some flexibility in the questions that students use based on their specific standard course. The second recommendation builds on the first recommendation but replaces student self-evaluation questions with a series of questions that address the student's experience in the virtual realm.

Recommendation 1: For standard (i.e., face-to-face) classes, a set of questions should be common across all courses within categories that allow for equitable evaluation regardless of class size or level. Within each category, instructors will be given the choice of choosing a default question or choosing alternate questions that are tailored for their particular course. All questions will be pre-approved by a faculty committee to ensure consistency and accuracy between instructors and courses. The course evaluation question bank will be administered and maintained by the UCR Center for Teaching and Learning.

Background: UCR has employed one student evaluation form for over a decade and the current form (iEval) does not differentiate between class size or level. Allowing more flexibility in the evaluation instrument provides instructors more meaningful feedback to improve their teaching while allowing evaluators (peers, departments, colleges, etc.) an opportunity to assess the overall quality of instruction in different courses.

UC Berkeley employs a course evaluation question bank that gives instructors flexibility in designing their own evaluations that can be tailored to specific courses and formats (<https://teaching.berkeley.edu/course-evaluations-question-bank>). The question bank is updated by faculty members as course formats evolve and are maintained by the Berkeley Center for Teaching & Learning. Although instructors can choose from a bank of questions, the evaluation will always have four different categories that must be addressed through the evaluation process:

1. Instructor-specific question themes
2. Course-specific question themes
3. Student self-evaluation questions
4. Open-ended questions

Each category has a number of default questions that were based on faculty feedback, but faculty are free to use other questions that are available in a course evaluation question bank. The default questions for the first two categories (required) are:

CATEGORY 1: Instructor-specific question themes

- Presentation of Content: The instructor presented content in an organized manner
- Clarity of Expectations or Directions: The instructor explained concepts clearly
- Helpfulness/Availability: The instructor was helpful when I had difficulties or questions
- Useful/clear feedback on performance: The instructor provided clear constructive feedback
- Encouraging of participation/discussion: The instructor encouraged student questions and participation
- Overall Teaching Effectiveness (MANDATORY): Considering both the limitations and possibilities of the subject matter and the course, how would you rate the overall effectiveness of this instructor?

CATEGORY 2: Course-specific question themes

- Course Content: The course was effectively organized
- Application & Skill Development: The course developed my abilities and skills for the subject
- Theory/Content Knowledge: The course developed my ability to think critically about the subject
- Course Overall: Considering both the limitations and possibilities of the subject matter and course, how would you rate the overall effectiveness of this course?

CATEGORY 3: Student Self-Evaluation Questions

1. How many class sessions did you attend?
2. On average, how many hours per week have you spent on this course (or section), including attending classes, doing readings, reviewing notes, writing papers, and any other course-related work?
3. How satisfied were you with your effort in this course?

CATEGORY 4: Open-Ended Questions

1. Please identify what you consider to be the strengths of the course.
2. Please identify area where you think the course (or section) could be improved.
3. Feedback for other students: What advice would you give to another student who is considering taking this course?

Recommendation 2: For non-standard courses (online, hybrid, flipped) that require a virtual interface, the same form that is used for standard courses can be used for categories 1, 2, and 4. Questions in category 3 should be replaced to reflect those aspects of the course that directly impact student learning and the use of technology and degree of engagement in the absence of personal contact.

Background: Many institutions of higher education are still struggling with how to effectively evaluate online formats for courses. Some professional organizations that offer online courses have tackled this issue and, in addition to questions that address overall instructor and course effectiveness, ask students a number of questions that are unique to online and hybrid courses. The American Historical Association is one of those organizations and has students answer the following questions in a course on Native American History (<https://www.historians.org/teaching-and-learning/teaching-resources-for-historians/teaching-and-learning-in-the-digital-age/the-history-of-the-americas/online-course-in-american-indian-history/student-evaluation-of-online-courses>):

1. Course-related resources were available for me
2. I was successful in this course as I am in similar courses taught in the classroom
3. I participated in this course as much as I would for similar courses taught in the classroom
4. I felt that I was a part of a learning community
5. I found it easy to remain motivated to complete the course requirements
6. The technology interfered with my learning
7. After taking this class, I am more comfortable using technology
8. Technical help was available when I needed it.
9. I would take another online course.
10. Prior to the start of this class I had completed:
 - No web-based classes
 - 1-2 web-based classes
 - 3-5 web-based classes
 - *More than 5 web-based classes*

Items 4 & 5: Subcommittee comments on a comprehensive model of teaching effectiveness, and how to implement the instructions (APM 210-1-d) to Review Committees when considering the teaching record of the candidate's performance.

Notes:

- Teaching effectiveness does not refer only to in-class teaching but includes all other aspects of educating students such as, but not limited to, general guidance, mentoring, supervising, and advising at all levels.
- The document pertains only to evaluation of teaching effectiveness by reviewing committees, not as a tool for continuous teaching improvement.

The sections of APM 210-1-d and the CALL that are relevant to these issues are included in Appendix 1.

APM 210-1-d spells out what evaluation committees should consider when judging teaching effectiveness (comprehensive model: e.g., command of the subject; ability to organize material and to present it with force and logic; capacity to awaken in students an awareness of the relationship of the subject to other fields of knowledge; fostering of student independence and capability to reason, etc.). Further, APM 210-1-d lists the most common types of evidence of teaching effectiveness that can be used by the evaluation committees (e.g., opinions of other faculty members; through student evaluations (SET); development of new and effective techniques of instruction, etc.). APM 210-1-d emphasizes that more than one kind of evidence should accompany each review file and that the committee should judge the total performance of the candidate with proper reference to assigned teaching responsibilities.

Our observation is that at UCR, for the most part, evaluation committees rely disproportionately (and often exclusively) on SET in evaluating teaching effectiveness, contradicting the requirement and the spirit of APM 210-1-d. An informal survey of the practices at other UC campuses reveals similar state (see Appendix 2). This is a problematic practice because recent research suggests that SET may not accurately reflect the quality of teaching due to biases and other factors (see for example, [here](#), [here](#), [here](#) and [here](#)), thus the need not to overemphasize SETs and certainly not to use them as the sole source for teaching evaluation.

Our recommendations below are based on secondary information gathering of practices at other universities (Appendix 3) and our own understanding and insights.

Key recommendation

In accordance with APM 210-1-d, the University should require that more than one kind of evidence of teaching effectiveness should accompany each review file. To implement this recommendation the following steps should be taken:

Education

The AP 210-1-d requirement should be widely disseminated to all faculty. It should be emphasized that it is the responsibility of the candidate to provide more than one kind of evidence of teaching effectiveness. The requirement should be emphasized in all orientation sessions to new faculty, department chairs, and CAP members.

Subcommittee B report for Ad Hoc Committee on Evaluation of Teaching

The Academic Personnel Office's current [Department Chair Checklist](#), Career Development section, item 1 states that department chairs should "Have annual reminder conversations with all faculty about The CALL and APM-210." As part of this conversation, we recommend that each department discuss what is the most appropriate evidence of teaching effectiveness (other than SET) in their own discipline/department and encourage faculty to provide such evidence when preparing their own evaluation file. A "one size fits all" notion of good teaching is not possible and, as such, disciplinary variations of the dimensions of effective teaching and the type of corroborating evidence are expected.

In support of this effort, the Center for Teaching & Learning should prepare a document (available online) that outlines common types of evidence of teaching effectiveness that are not SET. See, for example, the following sites:

[UCLA](#)

[The University of Kansas](#)

[University of Washington](#)

[University of Michigan](#) and [here](#)

The Center for Teaching & Learning in preparing the document on the common types of evidence of teaching effectiveness that are not SET should be informed by the discussion in the various departments.

We also recommend that the "Teaching Statements" in the eFile (the Department's Teaching Statement) should include a statement listing the most appropriate evidence of teaching effectiveness (other than SET) in their own discipline.

Taken together, action on these recommendations can provide faculty with much needed clarity about how teaching effectiveness is evaluated and the range of evidence that can be brought to bear on it.

Facilitation

To facilitate the process and to encourage candidates to provide multiple kinds of evidence of teaching effectiveness, the eFile collection system should be expanded to include additional explicit categories (under the Teaching Information) such as "peer-review", "lecture notes", "evaluation/assessment instruments", "measures of student achievement," etc. These categories should include an option to upload relevant documents. Further, the input category for "Syllabus" that is currently under "Other Information" should be moved to under the "Teaching Information" category.

For a specific recommendation on revising the structure of the eFile see Appendix 4.

"Enforcement"

Possibly, after transition period, after education and input facilitation would make providing more than one kind of evidence of teaching effectiveness routine, the eFile will not be considered complete unless the candidate provides more than one kind of evidence. At this point, CAP should return any departmental letter that does not discuss more than one kind of evidence of teaching effectiveness. We recommend that the "enforcement" stage would not commence without further endorsement by the Senate.

Members of the Subcommittee B Task Force

John Stamp, Subcommittee B Chair, University Writing Program

Jean Helwege, Professor, Area of Finance

John Levin, Distinguished Professor, Graduate School of Education

Rami Zwick, Professor, Area of Marketing

Jennifer Brown (ex-officio), Vice Provost for Undergraduate Education

Rich Cardullo (ex-officio), Professor, Evolution, Ecology, and Organismal Biology

Daniel Jeske (ex-officio), Vice Provost of Academic Personnel

Mariam Lam (ex-officio), Vice Chancellor of Diversity, Equity, and Inclusion

Ameae Walker (ex-officio), Vice Provost of Academic Personnel

APPENDIX 1 – APM 210-1 and the relevant sections from the CALL

APM 210-1

210-1. Instructions to Review Committees That Advise on Actions Concerning Appointees in the Professor and Corresponding Series

.....

210-1d. Criteria for Appointment, Promotion, and Appraisal

210-1d1. Teaching - Clearly demonstrated evidence of high quality in teaching is an essential criterion for appointment, advancement, or promotion. Under no circumstances will a tenure commitment be made unless there is clear documentation of ability and diligence in the teaching role.

In judging the effectiveness of a candidate's teaching, the committee should consider such points as the following (*bullet points are not in the original*)

- the candidate's command of the subject
- continuous growth in the subject field
- ability to organize material and to present it with force and logic
- capacity to awaken in students an awareness of the relationship of the subject to other fields of knowledge
- fostering of student independence and capability to reason
- spirit and enthusiasm which vitalize the candidate's learning and teaching
- ability to arouse curiosity in beginning students, to encourage high standards, and to stimulate advanced students to creative work
- personal attributes as they affect teaching and students
- extent and skill of the candidate's participation in the general guidance, mentoring, and advising of students
- effectiveness in creating an academic environment that is open and encouraging to all students, including development of particularly effective strategies for the educational advancement of students in various underrepresented groups

The committee should pay due attention to the variety of demands placed on instructors by the types of teaching called for in various disciplines and at various levels, and should judge the total performance of the candidate with proper reference to assigned teaching responsibilities.

The committee should clearly indicate the sources of evidence on which its appraisal of teaching competence has been based. In those exceptional cases when no such evidence is available, the candidate's potentialities as a teacher may be indicated in closely analogous activities.

In preparing its recommendation, the review committee should keep in mind that a redacted copy of its report may be an important means of informing the candidate of the evaluation of his or her teaching and of the basis for that evaluation.

Subcommittee B report for Ad Hoc Committee on Evaluation of Teaching

It is the responsibility of the department chair to submit meaningful statements, accompanied by evidence, of the candidate's teaching effectiveness at lower-division, upper-division, and graduate levels of instruction.

More than one kind of evidence shall accompany each review file.

Among significant types of evidence of teaching effectiveness are the following:

- (a) opinions of other faculty members knowledgeable in the candidate's field, particularly if based on class visitations, on attendance at public lectures or lectures before professional societies given by the candidate, or on the performance of students in courses taught by the candidate that are prerequisite to those of the informant;
- (b) opinions of students;
- (c) opinions of graduates who have achieved notable professional success since leaving the University;
- (d) number and caliber of students guided in research by the candidate and of those attracted to the campus by the candidate's reputation as a teacher; and
- (e) development of new and effective techniques of instruction, including techniques that meet the needs of students from groups that are underrepresented in the field of instruction.

All cases for advancement and promotion normally will include:

- (a) evaluations and comments solicited from students for most, if not all, courses taught since the candidate's last review;
- (b) a quarter-by-quarter or semester-by-semester enumeration of the number and types of courses and tutorials taught since the candidate's last review;
- (c) their level;
- (d) their enrollments;
- (e) the percentage of students represented by student course evaluations for each course;
- (f) brief explanations for abnormal course loads;
- (g) identification of any new courses taught or of old courses when there was substantial reorganization of approach or content;
- (h) notice of any awards or formal mentions for distinguished teaching;
- (i) when the faculty member under review wishes, a self-evaluation of his or her teaching;
- (j) evaluation by other faculty members of teaching effectiveness.

When any of the information specified in this paragraph is not provided, the department chair will include an explanation for that omission in the candidate's dossier.

If such information is not included with the letter of recommendation and its absence is not adequately accounted for, it is the review committee chair's responsibility to request it through the Chancellor.

Subcommittee B report for Ad Hoc Committee on Evaluation of Teaching

The CALL

.....

J. Departmental Recommendation Letter

J1b Evaluation of Teaching

Where possible and applicable, the departmental letter should comment on items such as the following:

- i. The role of the candidate in the graduate and undergraduate instructional program including such items as the amount relative to the department norms variety and difficulty of the teaching assignments and the preparation and attention given by the candidate to his/her teaching responsibilities. Make reference to teaching information form.
- ii. Out-of-class teaching and advising at both the graduate and undergraduate levels (careful thought should be given to the advising role of each candidate): directed research, special studies, help given to students, office hours with students, contributions to the teaching of other faculty, etc.
- iii. Graduate student supervision and advising: PhDs, Masters, committees, post-doctoral, and Graduate Research Assistant supervision. Attention may be given to the role of the candidate in attracting high caliber graduate students to the campus.
- iv. Development of new and effective techniques of instruction; writing of teaching materials, manuals, textbooks.
- v. Evaluation of teaching as judged by departmental colleagues. Guidelines dealing with the evaluation of teaching are contained in [APM 210-1](#) and should be consulted by Chairs on behalf of their departments. Among other elements of teaching, faculty colleagues are particularly well qualified to make thoughtful and substantial assessments of the candidate's command of subject matter and continuous growth in his/her field. Faculty perceptions derived from direct observation and information should be shared with colleagues at the departmental personnel meeting concerning the candidate and incorporated into the file in an appropriate manner.
- vi. Evaluation of teaching by students. Materials submitted by students should be discussed by the department in its meeting and summarized and evaluated in the departmental letter. All teaching evaluations performed during the review period should be assessed and commented on. Hearsay is not acceptable for use in teaching evaluations.

APPENDIX 2 - UC practices

April 18, 2019

Vice Provost for Academic Personnel (or Faculty) at the UC campuses were asked to comment on the following questions:

1. Implementing APM 210-d-1 that stipulate that there be more than one form of evaluation of teaching
2. Do you incentivize SET?
3. Do you use teaching portfolios to evaluate teaching?

The following is a summary of the replies:

UCSD

Our Center for Innovations in Teaching and Learning (CITL) has recently compiled a webpage of resources on documenting teaching effectiveness, which includes a guide for evidence for personnel reviews.

- 1) Everyone uses course evaluations, and almost everyone discusses graduate mentoring, and UCSD count those as two sources of evidence. A few departments have a culture of including syllabi, and a few get letters from students.
- 2) UCSD is struggling with incentivization, but are on hold worrying about it, as UCSD is likely to transition to a new online system very soon.
- 3) Seldom use teaching portfolio in a personnel file, but CITL has started to talk about them.

UC Davis

UC Davis requires peer review of teaching for the two promotion steps; at any other step it is optional and not often done. The peer review typically involves one or more letters familiar with the candidate's subject matter material and involves attending at least one lecture and reviewing course materials. Dossiers without these will be sent back to the departments.

Student evaluation is not incentivized except in the School of Veterinary Medicine, which has local control over its entire curriculum. Students cannot register for the next semester's courses until they complete their course and instructor evaluations from the previous semester. UC Davis recognizes the limitations of SET both in undergraduate and graduate courses.

UC Davis Dean for Undergraduate Education would like to try an experiment with having a small team of trained faculty within a college do more exhaustive teaching evaluations for all faculty. So far, no college is willing to participate in such an experiment.

Policy. The Step Plus system for merit and promotion enables rewards for instructors who make outstanding contributions to excellence in teaching. Also, **peer evaluation of teaching is necessary for promotion.**

<http://academicaffairs.ucdavis.edu/policies/step-plus/guidelines-for-advancement-senate.html>

UCSF

#1 – there is variability by department and type of teaching (e.g., large lecture in prof. schools; small graduate seminars; individual lab-based instruction). Most dossiers contain structured evaluations of teaching from on-line evaluations (primarily professional school instruction) and individual letter of support for teaching contributions. UCSF does not send back files that contain only one type of teaching evaluation.

#2 – some professional schools do not release student grades until a certain percentage response rate (to on-line evaluations) is reached.

#3 – a minority of our faculty use teaching portfolios (primarily those associated with our campus' Academy of Medical Educators; but others as well).

UCSB

The UCSB Senate also has a task force investigating the statistical course evaluations.

1) in regard to APM 210-d-1, Departments generally provide both the statistical evaluation and the written student comments (there is a trial in some Departments for online evaluations, but the return response is low)

2) UCSB does not use any incentives

3) Faculty often provide a self-assessment of their teaching performance.

UCSB would like to have better statistical analysis of SET (and especially to educate the students on what these evaluations really are).

UCLA

In regard to #1, UCLA typically use student evals (numbers and comments, typically comments are much more informative), but also letters and/or forms from peers and departments. Procedures however are not standardized and different departments and schools use different approaches.

For #2 UCLA experience the lower rates with on-line evals. Using class time helps increase percentages of students evaluating. UCLA new Center for Advancement of Teaching [CAT, used to be Office of Instructional Development (OID)] is experimenting

with new evaluation programs and procedures. Also some of UCLA departments are adding procedures to evaluate special programs for faculty to participate in to help improve their teaching which will substitute for peer evaluation. This is still in an experimental and trial phase.

Other Forms of Evaluating and Reflecting on Teaching

Center for Education Innovation and Learning in the Sciences

<https://ceils.ucla.edu/resources/teaching-guides/other-forms/>

Given the many concerns with focusing only on student ratings of instruction, what are other ways that you and your institution can assess teaching effectiveness and student learning?

The Association of American Universities (AAU) created a [matrix](#) of various campus strategies in the evaluation of faculty teaching (e.g., promotion and tenure). The following is from the AAU matrix at <https://www.aau.edu/sites/default/files/AAU-Files/STEM-Education-Initiative/P%26T-Matrix.pdf>

UC Merced

For 1., course evaluations ratings are still the 'meat and potatoes' of assessment.

For 2., no incentives right now.

For 3., at review time, faculty will occasionally submit a set of auxiliary materials, such as syllabi, lesson plans, new material they have developed, etc. Often these are individuals about to undergo tenure review, and need to show how they have addressed areas of weakness identified in the MCA.

In addition,

UC MERCED try to make teaching stand out as important during onboarding. Encourage mentors to discuss effective teaching practices with pre-tenured faculty. The Center for Excellence in Teaching (CETL) gives presentations to all new faculty. CETL can also help informally assess and address weaknesses, when a faculty member seeks this help. Some departments encourage weaker instructors to go CETL, and many do not. CETL was approached (last year?) to help develop a more robust formal assessment approach, but CETL is not there yet.

UC Berkeley

CAP or the VP for faculty will send back threshold cases that limit the assessment of teaching to just reporting the means of student evaluations.

UC Berkeley still use a mix of online and paper forms. Online response rates are poor; although they can be boosted somewhat if the instructor asks students to bring their laptops to the last class and do the online evaluations then. UC Berkeley encourages portfolios.

The ninth meeting of the Berkeley Division of the Academic Senate (DIVCO) was held on Monday, February 11, 2019. One item on the agenda was a discussion of new proposed policies for the evaluation of teaching

https://academic-senate.berkeley.edu/sites/default/files/the_fortnightly_february_12_2019.pdf

DIVCO discussed proposed revisions to the policy on teaching evaluation. These revisions were promoted in part by [Philip Stark and Richard Freishtat's research](#) on the merits of the existing evaluations in assessing learning outcomes and teaching effectiveness. The Committee on Teaching had proposed a teaching dossier several years ago, and a new set of questions were proposed by the Committee on Budget and Interdepartmental Relations (Budget Committee) last year, and circulated to other Senate committees for feedback and commentary. Following this process, Vice Provost for the Faculty Ben Hermalin formulated an official revised policy. While there was consensus that the revised policy is an improvement on the current policy (which dates to 2006), DIVCO members raised concerns about increased workload both for staff compiling and entering data into APBears and for department chairs conducting peer reviews of teaching.

The VP document is at "Teaching_Evaluation_Memorandum_v2.pdf"

UCI

UCI requires a second piece of evidence for teaching. This is enforced because they use an electronic review platform and the file does not move forward unless the second piece is uploaded.

After a pilot of accepting any form of second evidence (3 years ago) UCI provides the guidance as this URL: <http://dtei.uci.edu/second-piece-of-evidence-of-teaching-effectiveness/>

The majority of faculty submit a reflective teaching statement as their second piece of evidence and this has been quite useful not only for review but for spreading and highlighting innovative things lots of faculty are doing when discussed at each level of review. However, UCI do have some that submit teaching portfolios, peer evaluations, and a very small number that continue to just submit syllabi (which was common first year but these are generally not helpful and UCI has discouraged this practice in absence of accompanying reflective teaching statement).

Policy. Evaluation System for P/T. Making changes toward using a three-bucket system to evaluate accomplishments in research, teaching and service separately.

While published UC policy indicates that at least two types of evidence should support evaluation of teaching (see the Supplemental Material or visit www.ucop.edu/academic-personnel-programs/files/apm/apm-210.pdf, p. 5), in practice, student evaluations are often the only evidence used. **For the 2016 review cycle, UC Irvine has required**

Subcommittee B report for Ad Hoc Committee on Evaluation of Teaching

individuals to upload at least one additional type of evidence to evaluate teaching (e.g., reflective teaching self-statement, syllabus, peer evaluation, or measure of student achievement). This change is a first step toward conducting a more thorough evaluation of the contributions to teaching. It also broadens the discussion of teaching by everyone involved in the review process and thus has the potential to increase awareness of the innovative and effective teaching practices taking place on campus.

(from CBE—Life Sciences Education Vol. 16, No. 4 Essay, Aligning Practice to Policies: Changing the Culture to Recognize and Reward Teaching at Research Universities. Published Online :15 Mar 2018 <https://doi.org/10.1187/cbe.17-02-0032>)

APPENDIX 3 – Survey of some non-UC practices

NSF Award Search: Award#1725959 - Collaborative Research: Transforming the Evaluation of Teaching: A Study of Institutional Change to Advance STEM Undergraduate Education
2017-2022

https://www.nsf.gov/awardsearch/showAward?AWD_ID=1725959&HistoricalAwards=false

In this project, the **University of Kansas**, the **University of Colorado at Boulder** and the **University of Massachusetts Amherst** will approach change in ways best suited to their campus culture. Each, though, will implement an evaluation framework that is based on two decades of scholarship on scholarly teaching and its evaluation. The framework draws on multiple sources of evidence, including students, peers, instructors, and that speaks to multiple dimensions of teaching and learning. The work on each campus will center on the development and use of a teaching evaluation rubric that provides a richer, more complete view of teaching practice, and the evidence that speaks to it.

See also at <https://www.aau.edu/sites/default/files/Blind-Links/Teaching-Eval-AAU-Chairs-Oct2018.pdf>

The University of Kansas

KU developing new, more insightful method of evaluating teaching than traditional student forms

03/18/2019

<http://today.ku.edu/2019/02/28/ku-developing-new-more-insightful-method-evaluating-teaching-traditional-student-forms>

BENCHMARKS FOR TEACHING EFFECTIVENESS

<https://cte.ku.edu/sites/cte.ku.edu/files/docs/Branding/Benchmarks/BenchmarkswRubricwCCUpdated.pdf>

see at “KU BenchmarkswRubricwCCUpdated.pdf”

USC

Excellence in Teaching - USC Provost

March 27, 2018

<https://www.provost.usc.edu/excellence-in-teaching/>

Academic Senate Task Force on Teaching Evaluations

The Academic Senate convened a Faculty Evaluation Task Force in 2013. The resulting white paper was reviewed by the full Academic Senate in December, 2013. The final paper was sent to the Provost's office in February, 2014. Teaching is covered on pages 3 and 4.

Teaching Evaluations Update | Academic Senate | USC

<https://academicsenate.usc.edu/teaching-evaluations-update/>

Subcommittee B report for Ad Hoc Committee on Evaluation of Teaching

September 20, 2018

Recommendations of the Faculty Evaluation Task Force (2013) of the Academic Senate
https://docs.google.com/document/d/13YvYZpm_HWAul8LYB5vJhBJdLuSvcZiZQPN6bTGISyA/edit

APPENDIX V. EXCERPTS FROM 2013 UCAPT MANUAL (p. 14)

University of Washington | Seattle, WA

A guide to best practice for evaluating teaching | Center for Teaching and Learning
<https://www.washington.edu/teaching/teaching-resources/assessing-and-improving-teaching/evaluation/>

- Self-assessment
- Peer review
- Student evaluations

University of Michigan

Evaluation of Teaching | CRLT

<http://www.crlt.umich.edu/resources/evaluation-teaching>

Evaluation of teaching can have many purposes, including collecting feedback for teaching improvement, developing a portfolio for job applications, or gathering data as part of personnel decisions, such as reappointment or promotion and tenure. Most of the methods described below can be used for all of these functions. What follows are multiple methods for collecting information about instructors' activities, accomplishments, and effectiveness in teaching, in the classroom and beyond. While this list includes best practices for using student ratings, it also offers suggestions for ensuring that student ratings are not the only source of evidence used to assess instructional effectiveness, an approach consistent with research. In addition, detailed resources are available on the topics of [student ratings of instruction](#), [peer review of teaching](#), and [teaching portfolios and course portfolios](#).

Guidelines for Evaluating Teaching | CRLT

<http://www.crlt.umich.edu/tstrategies/guidelines>

University of Oregon

Revising UO's Teaching Evaluations | Office of the Provost

<https://provost.uoregon.edu/revising-uos-teaching-evaluations>

Revising UO's Teaching Evaluations

The Office of the Provost and the University Senate are currently working together to critique and revise our entire teaching evaluation system. Recent research suggests that student ratings may not accurately reflect the quality of teaching due to biases and other factors [3, 4, 5, 6]. The University of Oregon's own assessment of student course evaluation ratings have corroborated these findings [2].

Subcommittee B report for Ad Hoc Committee on Evaluation of Teaching

The Association of American Universities (AAU) and other universities around the globe from University of Colorado, Boulder to University College London, England have argued that it is time for universities' practices regarding teaching excellence and evaluation to align with their policies [1, 7, 8]. As such, the University of Oregon seeks to develop a holistic new teaching evaluation system that does more than simply replace problematic evaluation instruments so that we can help the UO community more effectively define, develop, evaluate, and reward teaching excellence. More specifically, we want to increase equity and transparency in the use and efficacy of teaching evaluations for merit, contract renewal, promotion and tenure, while simultaneously providing instructors with tools for continual course improvement.

Draft April 29 2018

Prepared by Senate Task Force for Teaching Evaluation

Proposed UO Peer Review of Teaching Framework

https://provost.uoregon.edu/files/proposed_peer_review_framework_aug2018.pdf

The University of Texas at Austin

Peer Observation of Teaching | Faculty Innovation Center

<https://facultyinnovate.utexas.edu/peer-observation>

University of Wisconsin System

Student Evaluation of Instruction | Board of Regents

<https://www.wisconsin.edu/regents/policies/student-evaluation-of-instruction/>

In assessing teaching effectiveness, the Regents believe that student evaluations are an important and useful source of evidence that should be explicitly considered in reaching judgments. The Regents assert, **however, that student evaluation must not be a substitute for direct peer judgment of teaching effectiveness** through a variety of means such as observation of teaching, assessment of syllabi, examinations and other course materials, and evaluation of contributions to development and strengthening of departmental curriculum. Moreover, effective peer judgment of teaching effectiveness necessarily includes both examination of the faculty member's current level of performance, and also his or her potential for growth.

Purdue University

Task force to recommend alternatives for evaluating teaching

November 29, 2017

<https://www.purdue.edu/newsroom/purduetoday/releases/2017/Q4/task-force-to-recommend-alternatives-for-evaluating-teaching.html>

Boston University

Report and Findings of the Task Force on Evaluating Teaching » Office of the Provost
February 7th, 2019

<https://www.bu.edu/provost/2019/02/07/report-and-findings-of-the-task-force-on-evaluating-teaching/>

Report of the Task Force on Evaluating Teaching
October 2018

<https://www.bu.edu/provost/files/2019/02/Report-of-the-Task-Force-on-Evaluating-Teaching.pdf>

Executive Summary

In May 2017, the Task Force on Evaluating Teaching was charged with gathering information about the language and content of student course evaluations at Boston University and considering how this content aligns with current research and national conversations about how best to evaluate university teaching. The Task Force was further charged with drafting a University-wide course feedback form with both common content and questions and flexible options for schools, colleges, programs and departments to tailor the surveys to their specific needs. **Additionally, the Task Force was asked to make recommendations about a move to collect student course feedback online. Finally, the Task Force was invited to comment on the place of student course feedback in the overall process of evaluating teaching on campus, considering the role of other measures, such as peer evaluation of teaching, self-reflection, and review of portfolios of teaching materials, all with the goal of creating a set of standard guidelines and practices for evaluating teaching on the Charles River and Medical Campuses.**

Key recommendations

3. The University should require multiple measures in order to responsibly evaluate teaching, including peer evaluation and teaching portfolios, for all personnel decisions, including contract renewal.

APPENDIX 4 – Proposed changes to the eFile structure

eFile Categories

CV Overview

- Publications
- Creative Activities
- Patents
- Professional Service
- University and Public Service
- **Teaching Information**
 - Teaching Information and Evidence of Teaching Effectiveness
 - Teaching Records and Student Evaluation of Teaching
 - Teaching Releases
 - Departmental Teaching Statement
 - Faculty's Self Statement of Teaching Objectives and Performance
 - Pedagogical Activities (*allow upload of documents*)
 - Enter "Description"
 - Peer-Review of Teaching Effectiveness
 - Curriculum Developments
 - Program Development
 - Course Development
 - Course Revision
 - Development of New and Effective Techniques of Instruction
 - Representative Teaching Materials
 - Syllabus
 - Teaching Notes
 - Examples for Assessment Used
 - Videos of Lectures
 - Other
 - Letters From Former Students Who Have Achieved Notable Professional Success (Non-Confidential)
 - Extension Teaching
 - Mentoring
 - Advising
 - Post Doc
 - Summer Internship
 - Senior Thesis Advisor
 - Mentor-No UCR Student
 - Dissertation Reading-Non UCR Student
 - Visiting Researcher Visiting Scholar
 - Thesis Director Non UCR Student
 - Non UC-Teaching
 - Other

The category of “Student Instruction and Sponsorship” should not be a new category but should be included under the general category of “Pedagogical Activities”.

- **Student Instruction and Sponsorship**

- Student Instruction
 - Committee (required)
 - PhD Dissertation
 - Qualifying Exam
 - Prospectus
 - Masters Thesis
 - Oral Exam
 - Advisory Committee
 - Analytical Project
 - Role
 - Major Professor
 - Co-Major Professor
 - Chair
 - Co-Chair
 - Member
 - Thesis Advisor
 - Other
- Student Sponsorship
 - Type of Support (required)
 - Training And Grant
 - Student Grant
 - Non-UC Fellowship
 - UC Fellowship
 - Research Grant
 - Intramural Funds
 - In Kind Sponsorship
 - Other
- Fellowship, Grant, and Gift Activities
- Memberships
- Honors and Awards
- Employment History
- Education
- Self-Statements
- Other Information
 - Non-Confidential Document (Type of document - required)
 - Comments from Extramural Grant Reviewers Contract
 - Invitation Letters
 - News Articles
 - Public Acceptance Letter Reader's Reports
 - Thank-you Letter
 - Unsolicited Letter (Non-Confidential)
 - Other
 - Letter from Other Departments/Programs, Institutes and Centers (Proposed action – required)