Faculty Evaluation Research & Recommendations
Senate Diversity Subcommittee, Sonoma State University
SDS Chair: Lisel Murdock-Perriera
Task Force Members & Authors: Sil Machado, Teresa Nguyen, Patrick Johnson, Aidong Hu, Krista Altaker

Introduction

In higher education, review of a faculty's teaching effectiveness is one of the primary means to ensure the quality of teaching. Teaching evaluations can explicitly shape careers, disciplines, and the entire university.

Bias in teaching evaluations has also been explicitly recognized in the most recent California Faculty Association (CFA) collective bargaining agreement (CBA). In the latest news from CBA update, CFA stated:

> Faculty now have the explicit right to address and rebut bias in student evaluations. Also, reviews must consider the rebuttals when reviewing the underlying student opinions. Student opinion surveys play an important role in reviews of faculty, and underline research shows that women, LGBTQIA+, and BIPOC faculty consistently receive unfair and unhelpful feedback that reveals bias on the part of evaluators.

In order to address bias in the teaching evaluation process, an examination and revision of current approaches is essential. By considering faculty's diverse backgrounds to generate meaningful feedback, evaluation reform may contribute to career development and enhance the success of faculty's teaching and scholarship activities.

<u>**Review of Evidence: Bias in Faculty Evaluation**</u>

Drawing on 80 years of student ratings research, Linse (2017) found that, historically, most research on bias in student evaluations focuses on gender. However, more recent research, cited below, demonstrates that bias in student evaluations also centers on faculty race and ethnicity. Linse argues that student ratings are often misinterpreted by administrators and faculty evaluators and reiterates that student ratings are student perception data and do not measure what they learned in a course.

- **Racial Bias in SETEs.** Racial bias in student evaluations of teaching effectiveness (SETEs) is prevalent and well documented. Numerous studies have shown that Black faculty and faculty of color receive lower SETE scores than white faculty. SETEs provide data that reflect the demographics of the students and their perceptions of their teacher and/or class learning experience. **Aversive Racism**, "a pervasive, unconscious prejudice that occurs even among well-intentioned people" (Aruguete, Slater & Mwaikinda, 2017, p. 499), has been identified as a key aspect of bias present in student evaluations. Demographics of students and other unrelated factors impact SETE scores more than actual teaching or course design (Heffernan, 2021, Chavez & Mitchell, 2020, Steinberg and Sartain, 2020). Faculty of color and other underrepresented minority faculty are the primary recipients of negative and abusive comments from SETEs (Heffernan, 2021). Two sources reflect on **legal implications** of known bias in SETEs,

identifying potential for cases of defamation and breach of duty to care (Reinsch, Goltz & Hietapelto, 2020; Heffernan, 2021).

- **Racial bias in Scholarship.** Racial bias in scholarship reflects another area of inequity in academic career processes. Black faculty reflect that they "have to be twice as good…publish twice as high, and do more of it" (Griffin et al., 2013, p. 504). Many are still held to the "publish or perish" model, reflecting "white logic" and "white methods" (Spalter-Roth et al., 2019); with hidden criteria in place under what appears to be a "fair" and equitable process (Griffen et al, 2013). Turner and colleagues (2008) report that faculty of color often report a love of teaching, but undervaluation for their scholarly work, and that challenges to credentials and intellect in the classroom lead to dissatisfaction. Faculty of color often experience "illegitimization of their research and methodologies in academic culture, scholarly journals, disciplinary associations, professional networks and funding entities" (Turner et al., 2008, p. 146).

- **Racial Bias in Service.** Faculty of color are more likely to engage in liaison-related service, and to be involved as leaders in professional organizations, than their white counterparts (Wood et al., 2015). Many Black female faculty report not being supported in making the choice to turn down service opportunities to prioritize scholarship, teaching or simply seek balance (Griffen et al., 2013). Further, the experience of being tagged as the sole representative for their race and the ensuing isolation as a minoritized faculty member exacerbates the already marginalized experience for faculty of color (Turner et al. 2008). Domingo et al. (2020) report the inequitable distribution of academic service with no reward, devaluing of service by the university, and lack of clarity of the role of service in the RTP process. All of these intersect with other barriers relating to racist stereotypes and gender bias (described below).

- **Gender Bias in SETE.** For their study on biases in student evaluations, Chavez and Mitchell (2020) presented students with welcome videos for identical online courses. These videos were students' only exposure to the professors' gender and race or ethnicity. Controlling for other course factors, Chavez and Mitchell (2020) found that female instructors of color scored lower on ordinal student evaluations and received more negative comments than white male instructors. The authors point to experiences as evidence of the systemic problems that lead to higher attrition within university settings. Heffernan (2017) found that women faculty of all races experienced significant bias in student evaluations, finding that they received thirty-seven percent lower than men. Mitchell and Martin (2018) note that gender bias informs the substance of student evaluations. They observe that students are more likely to comment on women faculty's personality and appearance. They also noted that students were more likely to refer to women faculty as "teacher," (versus "professor" or "doctor") potentially denoting a lack of professional respect for women faculty. Ultimately, they assert, "Students appear to evaluate women poorly simply because they are women" (p. 5).

- **Gender Bias in Service.** Domingo et al (2020) found that imbalanced service workloads served as a significant impediment in the career advancement of women faculty, specifically women of color. Looking at the experiences of mid-career women faculty, Hart (2016) found that service expectations compromised women faculty's ability to excel in other, more highly valued, aspects of being a professor, such as teaching and research. . Relatedly, Guarino and Borden (2017) suggest that women faculty's focus on service potentially impacts their productivity in research and teaching, activities that they note "can lead directly to salary differentials and overall success in academia" (p. 690).

- **Gender Bias in Overall RTP.** Weisshaar (2017) finds that there is a gender gap in terms of promotion to tenure. Drawing on the career histories of former assistant professors in sociology, computer science, and English, Weisshaar found that women were less likely to gain tenure in the department where they started their career and on average earned tenure in "less prestigious departments" than men in their fields. Weisshaar finds that the gender gap was not explained by productivity (or lack thereof). Rather, Weisshaar argues that the inequality is rooted in the RTP evaluation process.

- **Sexual Orientation Bias.** Sexual orientation bias has been identified in contemporary literature (Heffernan, 2021). A study by Russ, Simond, and Hunt (2008) correlated student perception of homosexual instructors with student evaluations of the instructors as less competent, and less credible. Students in the study reported adecreased sense that they could learn as much from non-heterosexual instructors as compared to heterosexual instructors.

Overall RTP Impact of Bias - Intersectionality
The synthesis of evidence for racial and gender bias, and bias toward other under-represented faculty in various forms contributes to an increasing understanding of the overall impact of bias across the RTP process as a whole. A few key concepts that arose out of the literature include, but are not limited to:

- **Racism Cost**: In their narrative research study exploring the experiences of faculty of color at predominantly white institutions, Arnold, Osanloo, and Newcomb (2021) found that faculty of color were particularly vulnerable to what they term as a "retaliation tax" for their justice-oriented work. Respondents described taking extra precautions to assuage white students' potential discomfort with justice-oriented course material prior to the start of course. Despite their best efforts, respondents detailed the ways that students retaliated against through their student evaluations. As one respondent put it, "My job is to provide space to discuss difficult issues in leadership and policy. As a result of dealing with difficult students in my class, I was reported to the campus diversity office for being racist!" (p. 132). Such experiences are representative of a larger "Black Tax" that is levied on Black faculty, manifesting in encounters with the institution that are emotionally damaging and contribute to emotional and physical exhaustion. Such experiences are part of the "high cost to pay" to be a Black faculty member. (Heffernan, 2021, Steinberg & Sartain, 2020).

- **Scholarship Topics of Race/Ethnicity and Feminism Devalued**: Roos and Gatta (2009) noted that institutions devalued the feminist scholarship produced by women. One respondent stated that their initial tenure case was rejected because of their subject matter, citing a lack of support from their dean and department. In comparing the reception of feminist scholarship from women to "quality of mind" scholarship produced by men, one respondent stated, "As a feminist woman you have to work twice as hard just to get equitable promotion decisions" (p. 188). The biases related to feminist scholarship and that produced by women, more broadly, can also manifest in open hostility towards the scholar and their work, animus that comes to bare in the tenure promotion process. Looking at the barriers to promotion for women law professors of color, Deo (2017) observes that these scholars often publish articles and essays that explore the intersection of the law with race, gender, sexual orientation and other forms of identity. Deo contends that such work "may be viewed with suspicion by colleagues who, at best, do not understand them, but at worst feel threatened by the work itself — and respond by derailing a junior colleague's tenure or promotion application" (p. 1024).

Recommendations to Reduce Bias in Faculty Evaluation & Foster Faculty Equity
The following recommendations, drawn from research on faculty evaluation, are offered to FSAC in the hopes that they can be included in the guidelines provided to faculty evaluation committees.

- **Multiple Forms of Evaluation:** Departments should utilize multiple forms of evaluation for faculty teaching, including learning outcome measures, alumni ratings, instructor self-assessments, and faculty portfolios.

- **Use Robust Teaching Observations Practices & Policies:** Departments should develop robust teaching observation policies and procedures. Practices to consider include:
    - A thorough review of the syllabus and course materials, including texts, online learning materials utilized, and assignments for the course. Evaluating multiple dimensions of teaching, including those that occur outside the classroom, allows for a richer understanding of the faculty's pedagogical choices and how they align with course learning objectives.
    - A faculty interview prior to the observation during which the observer and faculty review information about the course, the faculty's pedagogical approach in the course and its evolution over time, the course objectives the faculty intends to address during the observation and how student learning will be assessed, and targeted feedback the faculty would like to receive from the observation.
    - A faculty interview following the observation during which the faculty can provide further context for the decisions made during the observation.
    - Evaluations must be sent to the candidate within 10 days per University policy. Candidates sometimes find themselves in the uncomfortable position of having to remind senior faculty about this. We recommend that FSAC encourage

departments to include the 10-day policy on their peer observation paperwork so this is more closely tracked by those conducting teaching observations.

- ○ Departments and Schools should consider other forms of peer observation emphasizing mutual collaboration. For example, faculty could participate in "Teaching Triangles" composed of three faculty in related or unrelated fields. These three peers would rotate observation and feedback responsibilities.

- **Use Robust Procedures for Interpreting SETE's:** Evaluators should have clearly defined process for interpreting SETE information to counter the known flaws in this data (both in source and methodology). Awareness of the impact of student bias and the disproportionate effect on faculty from marginalized groups can help to dismantle these tools of systemic oppression. *Departments should develop a robust procedure for interpreting SETEs, including:*
    - ○ Evaluators should be trained on how to appropriately interpret SETEs. Training can include workshops or handouts that summarize the findings of Linse (2017).
    - ○ SETEs should not be used to compare a candidate to another faculty member or to departmental averages.
    - ○ Evaluators are encouraged to include text describing the biased nature of SETEs. We include boilerplate language in this document.

- **Take a Longitudinal View of Teaching Effectiveness:** Recognize and acknowledge in faculty evaluations that a longitudinal view of teaching effectiveness should be taken, rather than focusing exclusively on current teaching practices. SETE information should be considered longitudinally, by examining trends over the course of a faculty career, rather than focus on a single datapoint

- **Monitor for Bias in Faculty Evaluations:** Recognize bias exists, that all faculty evaluators are vulnerable to bias in their evaluations, and that faculty from marginalized groups are harmed most by evaluator bias. Recognize that bias in faculty evaluations contributes to the problems that systematically exclude faculty from marginalized groups from academe in favor of faculty from dominant groups. *Evaluators should carefully monitor for bias in their evaluations of faculty, especially in the evaluations of faculty from marginalized groups, and document steps taken to reduce bias in their evaluations.*

- **Monitor Language & Tone:** Research has demonstrated that faculty evaluators or observers are more likely to use stronger and more positive language when describing the performance of men (e.g., "emerging star"). *Evaluators should carefully monitor any difference in language and tone when writing letters for candidates, and committee members should hold each other accountable and provide feedback to each other when differences in language and tone are detected.*

- **Advocate for Hazard Pay:** Recognizing the impact of abusive and harmful comments in SETE's (Heffernan, 2021, Steinberg & Sartain, 2020, Linse, 2017, Arnold et al., 2021), in their evaluations and RTP recommendations, department and school committees should

advocate for hazard pay and other compensation for faculty from underrepresented groups.

Suggested Language for RTP Committees at Department, School, and University Levels
Below we provide boilerplate text to include in your evaluation letter for faculty. Please consider adding some or all of the text below.

Statement Acknowledging Biased Nature of SETEs:
"We urge readers to give extreme caution when interpreting SETE results given the wide research base showing that sexism, racism, and homophobia influence student evaluations of teaching effectiveness (for review, see Heffernan, T., 2021). Specifically, research has shown that women, people of color, and members of the queer community are more likely to have low quantitative ratings on SETEs and more negative open-ended student comments. These biases are particularly exacerbated when faculty members teach difficult subject areas."

Statement Stressing the Importance of Other Measures of Teaching Effectiveness:
"Given these biases, our review of Professor _____'s teaching more appropriately relies on peer observations, self-assessment of teaching, and course materials (syllabi, assignments, etc.)."

<u>Recommendations to FSAC</u>
The following recommendations are offered to FSAC from SDS to support FSAC in fostering a more equitable and inclusive experience for faculty teaching at Sonoma State University.

- Given the clear and convincing evidence about bias in SETE's and the very real professional and psychological harm this causes to faculty from marginalized groups, FSAC should call for an immediate change of policy that outlines a plan to discontinue the use of SETE's at Sonoma State University. SETE's should be replaced with new methods of gathering data from students about their learning experiences that are less prone to student bias or for which bias can be detected. These may include:
    - Student focus groups
    - Student interviews
    - Pre- and post-class assessments of student learning outcomes
    - Professor-designed surveys aimed toward teaching effectiveness, including mid-semester feedback.
    - Teaching dossiers that include evidence of curricular engagement (e.g., examples of student work that shows student learning outcomes or a comparison of student work at the beginning and end of term; submitting rubrics; completed Teaching Practice Inventory, etc.)

    FSAC should identify new methods of assessment and work with departments to pilot them to support the University to develop new policies and procedures for assessing student learning experience.

- SDS recommends the discontinuation of the use of SETE's. In the interim, we suggest making immediate changes to existing SETE's:

- ○ The literature suggests that bias literacy training for students prior to providing feedback on their learning experiences in classes can help to mitigate bias against faculty from marginalized groups. *FSAC should develop a plan with the University to provide such training so that meaningful data can be collected from students about their learning experiences without the negative implications for marginalized faculty.*
- ○ Closed-ended and open-ended questions should explicitly prompt students to reflect on their learning experiences rather than ratings of teaching effectiveness as students are unqualified to evaluate an instructor's knowledge and competence. *The SETEs should be renamed to reflect this framing. For example, 'Student Learning Survey' would be more appropriate than 'Evaluation of Teaching Effectiveness.'*
- ○ *Faculty and evaluators should have the ability to view the results dynamically–that is, filtering out or grouping ratings based on various factors like students' expected grade, rating of the difficulty of the subject matter, the frequency of attendance, and the number of hours worked per week on the course.*
- ○ *FSAC should have a formal process to remove discriminatory, biased, or demeaning SETE comments and remove associated quantitative data from that student. Instructors should have the ability to initiate the process for removal multiple times per year.*
- ○ In support of faculty experimentation with innovative teaching techniques, *FSAC should develop a policy that allows faculty members to pick a semester every couple of years in which SETE ratings can be tossed out and not counted toward employment decisions.*

- ● Given the harm experienced by under-represented faculty due to SETE abuse and bias (Heffernan, 2021, Steinberg & Sartain, 2020, Linse, 2017, Arnold et al., 2021), consider recommending hazard pay or other forms of compensation to acknowledge the added challenges they face.

References

Arnold, N., Osanloo, A. F. & Newcomb, W. S. (2021). Paying Professional Taxes for Promotion and Tenure: The Costs of Justice Work for Black Faculty. *Journal of Research on Leadership Education 16(2), 122-139.* https://doi.org/10.1177/19427751211002220

Boren, J. P. & McPherson, M. B. (2018). Is Coming Out in the Classroom Still an Occupational Hazard? A Replication of Russ, Simonds, and Hunt. Communication Studies. 69(3), 242-250.

Chávez, K., & Mitchell, K. M. (2020). Exploring bias in student evaluations: Gender, race, and ethnicity. *PS: Political Science & Politics*, *53*(2), 270-274.

Deo, M. E. (2017). Intersectional barriers to tenure. *UCDL Rev.*, *51*, 997.

Domingo, C. R., Gerber, N. C., Harris, D., Mamo, L., Pasion, S. G., Rebanal, R. D., & Rosser, S. V. (2020). More service or more advancement: Institutional barriers to academic success for women and women of color faculty at a large public comprehensive minority-serving state university. *Journal of Diversity in Higher Education*.

Griffin, K. A., Bennett, J. C., & Harris, J. (2013). Marginalizing Merit?: Gender Differences in Black Faculty D/discourses on Tenure, Advancement, and Professional Success. *The Review of Higher Education 36(4), 489-512.* https://doi.org/10.1353/rhe.2013.0040

Guarino, C. M., & Borden, V. M. (2017). Faculty service loads and gender: Are women taking care of the academic family?. *Research in higher education*, *58*(6), 672-694.

Hart, J. (2016). Dissecting a gendered organization: Implications for career trajectories for mid-career faculty women in STEM. *The Journal of Higher Education*, *87*(5), 605-634.

Heffernan, T. (2022). Sexism, racism, prejudice, and bias: a literature review and synthesis of research surrounding student evaluations of courses and teaching. *Assessment & Evaluation in Higher Education*, *47*(1), 144-154.

Linse, A. R. (2017). Interpreting and using student ratings data: Guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation 54, 94-106.*

Mitchell, K. M., & Martin, J. (2018). Gender bias in student evaluations. *PS: Political Science & Politics*, *51*(3), 648-652.

Reinsch, R. W., Goltz, S. M., & Hietapelto, A. B. (2020). Student Evaluations and the Problem of Implicit Bias. *Journal of College and University Law, 45(1), 114–140*.

Roos, P. A. & Gatta M. L. (2009). Gender (in)equality in the academy: Subtle mechanisms and the production of inequality. *Research in Social Stratification and Mobility 27, 177-200.*

Russ, T. L., Simonds, C. J., & Hunt, S. K. (2002). Coming out in the Classroom…An Occupational Hazard?: The Influence of Sexual Orientation on Teacher Credibility and Perceived Student Learning. *Communication Education 51(3): 311-324.*

Spalter-Roth, R., Shin, J. H., Smith, J. A., Kalb, A. C., Moore, K. K., Cid-Martinez, I. & Toney, J. (2019). "Raced" Organizations and the Academic Success of Underrepresented Minority Faculty Members in Sociology. *Sociology of Race and Ethnicity 5(2), 261-277. DOI: 10.1177/2332649218807951*

Steinberg, M. P., & Sartain, L. (2020). What explains the race gap in teacher performance ratings? Evidence from Chicago Public Schools. *Educational Evaluation and Policy Analysis*. Prepublished December 10, 2020. http://doi.org/10.3102/0162373720970204

Turner, C. S. V., Gonzalez, J. C. & Wood, J. L. (2008). Faculty of Color in Academe: What 20 Years of Literature Tells Us. *Journal of Diversity in Higher Education 1(3) 139-168. DOI: 10.1037/a0012837*

Weisshaar, K. (2017). Publish and perish? An assessment of gender gaps in promotion to tenure in academia. *Social Forces*, *96*(2), 529-560.

Wood, J.L., Hilton, A. A. & Nevarez, C. (2015). Faculty of Color and White Faculty: An Analysis of Service in Colleges of Education in the Arizona Public University System. *The Journal of the Professoriate 8(1), 85-109.*