

Position statement on EPFL LEX 2.5.1 (Teaching Evaluations)  
Draft, 26.04.2022 SN/SP

Proposed modifications to LEX 2.5.1

Numerous scientific studies report that student evaluations of teaching reflect general biases related to gender, race, and ethnicity (see attached index). Experiences of female faculty at EPFL align with these findings, with numerous reports of low and sometimes derogatory student feedback, reflecting bias. In light of this, the expansion of the frequency and role of student evaluations of teaching at EPFL is deeply concerning.

Inclusion of the results of teaching evaluations in promotion and tenure proceedings means that career decisions are partially based on these biased metrics. At EPFL, currently, only in-depth student evaluations are included in dossiers for promotion and tenure. Teachers or section directors have the right to request an in-depth evaluation, or they should occur automatically once every 5 years. The change to conducting in-depth student evaluations of each course every semester will thus provoke a profound change to promotion and tenure proceedings, even without change to the language of the article on promotion and tenure. By extending in-depth evaluations to every course the amount of biased data available is drastically multiplied. Simply through the amount of data and the accessibility of it, there is a danger of its impact on promotion processes growing, unnoticed.

Moreover, discussion begun with the proposed changes to LEX 2.5.1 reveals a lack of broader official acknowledgement of bias in relation to teaching evaluations. Two sets of measures are thus proposed: specific changes to LEX 2.5.1 and additional changes that fall outside the purview of this directive.

*Proposed revisions to LEX 2.5.1*

#### **Article 5 Indicative feedback**

Moving the indicative feedback earlier in its current form (“Overall, I think this course is good.”) has the potential to increase bias in responses. After only a few weeks in the course, it would be difficult to assess course quality in any meaningful way. Indicative feedback should instead be used to spark discussions regarding whether students understand the course content, structure, and learning outcomes, thereby providing feedback that the instructor can respond to without judgments about course quality (good/bad). The question asked should be determined by the instructor, e.g.: What do you expect from this course?

#### **Article 6 In-depth evaluation**

Guidelines should be provided by CAPE to Section Directors on the composition of a questionnaire to mitigate bias (e.g. restricting or eliminating qualitative comments). However, it is important to be aware that even through a careful composition of the questionnaire, eliminating bias completely is not possible. Access to course evaluation results should be limited to Section Directors and Instructors, not to the Office of the Dean and not to students.

It would be sufficient to conduct once every 2-3 years, not every semester.

#### **Article 9 Evaluation of examinations**

The Section Director should be trained by bias and discrimination experts to account for bias in their comparative analysis of course evaluations. In the event of significant disparities, bias should be considered before measures are proposed to resolve unfavorable evaluations.

**Article 10 Teaching Portfolio**

Candidates should have the right to determine the contents of their dossier. Teaching evaluations should be included in the promotion/tenure dossier at the discretion of the candidate only. We suggest eliminating this article because promotion/tenure procedure should be delineated in Lex 4.2.1 and 4.2.2.

**Throughout**

Use inclusive language

*Broader Proposals*

**Evaluating teaching for Promotion and Tenure** Students evaluations should not be the only means of evaluating teaching as they give only a partial and biased view. Student evaluations should be complemented at the discretion of the candidate with syllabi, teaching material, or voluntary evaluations from CAPE or colleagues.

**Monitoring / accounting for bias** EPFL should first and foremost acknowledge the role of bias in teaching evaluations and develop a plan for accounting for and mitigating it and its impact. CAPE should analyze or support the analysis of course evaluation results to monitor bias.

**Addressing bias** Section Directors and CEA committee members should be trained so that they are aware of bias in teaching evaluations and consider it in their recommendations and decision making

**Access** Since the evaluations are biased, it is critical that EPFL exercise care in access to the evaluations.

Kreitzer, R.J., Sweet-Cushman, J. Evaluating Student Evaluations of Teaching: a Review of Measurement and Equity Bias in SETs and Recommendations for Ethical Reform. *J Acad Ethics* (2021). <https://doi.org/10.1007/s10805-021-09400-w>

*Across all the studies in our sample, the clearest evidence of gender bias is in qualitative comments. Scholars employing content analysis of qualitative comments finds clear evidence of bias with women faculty and faculty of color are more likely to receive negative comments about personality traits, appearance, mannerisms, competence, and professionalism compared to white men (Wallace et al., 2019). Furthermore, many faculty report particularly mean-spirited and cruel comments (Lindahl & Unger, 2010). Instead of asking for general “comments,” assessments should direct students to provide feedback on certain experiences with the course, as this may reduce irrelevant and mean comments.*

*Essentially, evaluations are shaped by discipline, student interest, class level, class difficulty, class meeting time, and other course-specific characteristics, but not generally actual instructor quality (Franklin & Theall, 1995; Greenwald & Gillmore, 1997; Miles & House, 2015; Spooren et al., 2013; Wigington et al., 1989; Uttl et al., 2017, 2013; Wachtel, 1998).*

*However, Rosen (2018), using a massive (n = 7,800,000) Rate My Professor sample, finds there is no discipline where women receive higher evaluative scores. These lower scores may also be substantial, with one study finding that, controlling for other factors, female instructors received average ratings that were one-half standard deviation lower than men’s ratings (Hamermesh & Parker, 2005).*

*Specific conditions where women are at a disadvantage over their male colleagues abound. Disparate research demonstrates that men are perceived as more accurate in their teaching, have higher levels of education, are less sexist, more enthusiastic, competent, organized, professional, effective, easier to understand, prompt in providing feedback, and are less-harshly penalized for being tough graders*

Adams, S., Bekker, S., Fan, Y. et al. Gender Bias in Student Evaluations of Teaching: ‘Punish[ing] Those Who Fail To Do Their Gender Right’. *High Educ* (2021). <https://doi.org/10.1007/s10734-021-00704-9>

*Based on qualitative analysis of almost 400,000 survey responses, we argue that the kinds of comments prevalent in SET show distinct and gendered differences both in terms of assessment of teaching practice and of expectations about behaviour and affect. We therefore conclude that SET measures conformity with gendered expectations about teachers, rather than teaching quality, in ways that disadvantage women in particular – actively ‘punish[ing] those who fail to do their gender right’ (Butler, 1999, p.178).*

Boring, Anne, and Arnaud Philippe. “Reducing Discrimination in the Field: Evidence from an Awareness Raising Intervention Targeting Gender Biases in Student Evaluations of Teaching.” *Journal of Public Economics* 193 (January 2021). <https://doi.org/10.1016/j.jpubeco.2020.104323>.

*The biases and stereotypes that drive discrimination are frequently unconscious, and thus individuals receiving anti-bias messages might think the campaign does not apply to them. Therefore, it is necessary to find successful interventions to make unconscious biases “conscious” and thus reduce discrimination against candidates.*

Mitchell, Kristina M. W., and Jonathan Martin. “Gender Bias in Student Evaluations.” *PS: Political Science & Politics* 51, no. 3 (2018): 648–52. doi:10.1017/S104909651800001X.

*Our analysis of comments in both formal student evaluations and informal online ratings indicates that students do evaluate their professors differently based on whether they are women or men. Students tend to comment on a woman's appearance and personality far more often than a man's. Women are referred to as "teacher" more often than men, which indicates that students generally may have less professional respect for their female professors. Based on our empirical evidence of online SETs, bias does not seem to be based solely (or even primarily) on teaching style or even grading patterns. Students appear to evaluate women poorly simply because they are women.*

*More important is the question of why this matters. Many universities, colleges, and programs use SETs to make decisions on hiring, firing, and tenure. Because SETs are systematically biased against women, using it in personnel decisions is discriminatory. In addition, this could have broader implications for women in all professional fields.*

MacNell, Lillian, Adam Driscoll, and Andrea N. Hunt. "What's in a Name: Exposing Gender Bias in Student Ratings of Teaching." *Innovative Higher Education* 40, no. 4 (August 1, 2015): 291–303.

*Assistant instructors led one online discussion group using their own name, and another group using the name of an assistant instructor of the opposite gender. Assistant instructors who used a female name and identity received significantly lower ratings than assistant instructors who used a male name and identity, regardless of their actual gender or teaching ability.*

*In sum, students evaluated instructors perceived as female more harshly than instructors perceived as male, demonstrating the existence of gender bias. The authors theorize that students expected instructors perceived as female to exhibit strong interpersonal traits, such as respectfulness and enthusiasm, and punished them for perceived failures to exhibit these traits. Conversely, instructors perceived as male were rewarded for going above-and-beyond when they exhibited these traits.*