

UC SANTA CRUZ

SELECT ANNOTATED BIBLIOGRAPHY ON BEST PRACTICES FOR EQUITY AND DIVERSITY IN HIRING – HIRING PROCESS

Compiled by the Advancing Faculty Diversity Workgroup, 2020-2021, UC Santa Cruz

Beware of redefining search criteria.

Uhlmann, E. L., & Cohen, G. L. (2005). Constructed criteria: Redefining merit to justify discrimination. *Psychological Science*, 16(6), 474-480.

Abstract. This article presents an account of job discrimination according to which people redefine merit in a manner congenial to the idiosyncratic credentials of individual applicants from desired groups. In three studies, participants assigned male and female applicants to gender-stereotypical jobs. However, they did not view male and female applicants as having different strengths and weaknesses. Instead, they redefined the criteria for success at the job as requiring the specific credentials that a candidate of the desired gender happened to have. Commitment to hiring criteria prior to disclosure of the applicant's gender eliminated discrimination, suggesting that bias in the construction of hiring criteria plays a causal role in discrimination.

Beware of the bias and facade of “fit.”

Damani K. White-Lewis (2020) The Facade of Fit in Faculty Search Processes, *The Journal of Higher Education*, 91(6), 833-857, DOI: [10.1080/00221546.2020.1775058](https://doi.org/10.1080/00221546.2020.1775058)

Abstract. Various concerns regarding the vitality and racial/ethnic composition of the academic profession have prompted new study of faculty search committees and hiring paradigms, most notably examining the term “fit” in candidate appraisals. Yet no study utilizes a candidate evaluation framework to investigate whether or not faculty members truly assess for fit, or if these assessments stifle diversification processes, especially in light of pervasive institutional efforts to reform faculty hiring. This study uses a critical person-environment fit framework and multiple case study methods to investigate how faculty search committee members individually evaluate and collectively select prospective early-career faculty. Results indicate that fit, as a system of assumptions, practices, and tactics designed to evaluate and select candidates based on organizational needs, was minimal in faculty searches. Instead, faculty relied heavily on idiosyncratic preferences to evaluate research, teaching, and service credentials, which also contained criterion that directly and indirectly averted diversity. Findings reveal how the review

and selection of candidates are as much, if not more, about individual committee preferences than organizational demands or congruence.

Use structured free recall to assess candidates.

Baltes, B. B., Bauer, C. B., & Frensch, P. A. (2007). Does a structured free recall intervention reduce the effect of stereotypes on performance ratings and by what cognitive mechanism? *Journal of Applied Psychology, 92*(1), 151–164.

Abstract. The purpose of this article was to extend previous work on the effect of racial biases on performance ratings. The 1st of 2 studies examined whether a structured free recall intervention decreased the influence of negative racial biases on the performance ratings of Black men. Results indicated that without the intervention, raters who endorsed a negative stereotype of Black men as managers evaluated Black men more negatively. However, the structured free recall intervention successfully reduced these effects. The second study examined in more detail the cognitive mechanisms underlying the success of the intervention. Results are consistent with the assumption that the reduction of the influence of racial biases under structured free recall conditions is a consequence of a modified strength threshold for retrieval of behaviors from memory.

Recognize that bias affects research impact, citation rates, author order, journal acceptances, grant awards, invited talks, and professional achievements.

Re. research impact:

Hofstra, B., Kulkarni, V. V., Galvez, S. M. N., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences 117*(17), 9284–9291.

Abstract. Prior work finds a diversity paradox: Diversity breeds innovation, yet underrepresented groups that diversify organizations have less successful careers within them. Does the diversity paradox hold for scientists as well? We study this by utilizing a near-complete population of ~1.2 million US doctoral recipients from 1977 to 2015 and following their careers into publishing and faculty positions. We use text analysis and machine learning to answer a series of questions: How do we detect scientific innovations? Are underrepresented groups more likely to generate scientific innovations? And are the innovations of underrepresented groups adopted and rewarded? Our analyses show that underrepresented groups produce higher rates of scientific novelty. However, their novel contributions are devalued and discounted: For example, novel contributions by gender and racial minorities are taken up by other scholars at lower rates than novel contributions by gender and racial majorities, and equally impactful contributions of gender and racial minorities

are less likely to result in successful scientific careers than for majority groups. These results suggest there may be unwarranted reproduction of stratification in academic careers that discounts diversity's role in innovation and partly explains the underrepresentation of some groups in academia.

Re. author order and citation rates:

Bendels, M. H., Müller, R., Brueggmann, D., & Groneberg, D. A. (2018). Gender disparities in high-quality research revealed by Nature Index journals. *PLoS one*, 13(1), e0189136.

Abstract. Background: The present study aims to elucidate the state of gender equality in high-quality research by analyzing the representation of female authorships in the last decade (from 2008 to 2016). Methods: Based on the Gendermetrics platform, 293,557 research articles from 54 journals listed in the Nature Index were considered covering the categories *Life Science, Multidisciplinary, Earth & Environmental* and *Chemistry*. The core method was the combined analysis of the proportion of female authorships and the female-to-male odds ratio for first, co- and last authorships. The distribution of prestigious authorships was measured by the *Prestige Index*. Results: 29.8% of all authorships and 33.1% of the first, 31.8% of the co- and 18.1% of the last authorships were held by women. The corresponding female-to-male odds ratio is 1.19 (CI: 1.18–1.20) for first, 1.35 (CI: 1.34–1.36) for co- and 0.47 (CI: 0.46–0.48) for last authorships. Women are underrepresented at prestigious authorships compared to men (Prestige Index = -0.42). The underrepresentation accentuates in highly competitive articles attracting the highest citation rates, namely, articles with many authors and articles that were published in highest-impact journals. More specifically, a large negative correlation between the 5-Year-Impact-Factor of a journal and the female representation at prestigious authorships was revealed ($r(52) = -.63, P < .001$). Women publish fewer articles compared to men (39.0% female authors are responsible for 29.8% of all authorships) and are underrepresented at productivity levels of more than 2 articles per author. Articles with female key authors are less frequently cited than articles with male key authors. The gender-specific differences in citation rates increase the more authors contribute to an article. Distinct differences at the journal, journal category, continent and country level were revealed. The prognosis for the next decades forecast a very slow harmonization of authorships odds between the two genders.

Re. author order:

Pico, T., Bierman, P., Doyle, K., & Richardson, S. (2020). First authorship gender gap in the geosciences. *Earth and Space Science*, 7, e2020EA001203.

Abstract. Although gender parity has been reached at the graduate level in the geosciences, women remain a minority in faculty positions. First authorship of peer-reviewed scholarship is a measure of academic success and is often used to project potential in the hiring process. Given the importance of first author publications for hiring and advancement, we sought to quantify whether

women are underrepresented as first authors relative to their representation in the field of geoscience. We compiled first author names across 13 leading geoscience journals from January 2013 to April 2019 (n = 35,183). Using a database of 216,286 names from 79 countries, across 89 languages, we classified the likely gender associated with each author's given (first) name. We also estimated the gender distribution of authors who publish using only initials, which may itself be a strategy employed by some women to preempt perceived (and actual) gender bias in the publication process. Female names represent 13–30% of all first authors in our database and are substantially underrepresented relative to the proportion of women in early career positions (30–50%). The proportion of female-name first authors varies substantially by subfield, reflecting variation in representation of women across geoscience subdisciplines. In geoscience, the quantification of this first authorship gender gap supports the hypothesis that the publication process—namely, achievement or allocation of first authorship—is biased by social factors, which may modulate career success of women in the sciences.

Re. author order and author number:

West, J. D., Jacquet, J., King, M. M., Correll, S. J., & Bergstrom, C. T. (2013). The role of gender in scholarly authorship. *PloS one*, 8(7), e66212.

Abstract. Gender disparities appear to be decreasing in academia according to a number of metrics, such as grant funding, hiring, acceptance at scholarly journals, and productivity, and it might be tempting to think that gender inequity will soon be a problem of the past. However, a large-scale analysis based on over eight million papers across the natural sciences, social sciences, and humanities reveals a number of understated and persistent ways in which gender inequities remain. For instance, even where raw publication counts seem to be equal between genders, close inspection reveals that, in certain fields, men predominate in the prestigious first and last author positions. Moreover, women are significantly underrepresented as authors of single-authored papers. Academics should be aware of the subtle ways that gender disparities can occur in scholarly authorship.

Re. journal acceptances:

Budden, A. E., Tregenza, T., Aarssen, L. W., Koricheva, J., Leimu, R., & Lortie, C. J. (2008). Double-blind review favours increased representation of female authors. *Trends in Ecology & Evolution*, 23(1), 4-6.

Abstract. Double-blind peer review, in which neither author nor reviewer identity are revealed, is rarely practised in ecology or evolution journals. However, in 2001, double-blind review was introduced by the journal *Behavioral Ecology*. Following this policy change, there was a significant increase in female first-authored papers, a pattern not observed in a very similar journal that provides reviewers with author information. No negative effects could be identified, suggesting that double-blind review should be considered by other journals.

Re. grant awards:

Erosheva, E. A., Grant, S., Chen, M. C., Lindner, M. D., Nakamura, R. K., & Lee, C. J. (2020). NIH peer review: Criterion scores completely account for racial disparities in overall impact scores. *Science Advances*, 6(23), eaaz4868.

Abstract. Previous research has found that funding disparities are driven by applications' final impact scores and that only a portion of the black/white funding gap can be explained by bibliometrics and topic choice. Using National Institutes of Health R01 applications for council years 2014–2016, we examine assigned reviewers' preliminary overall impact and criterion scores to evaluate whether racial disparities in impact scores can be explained by application and applicant characteristics. We hypothesize that differences in commensuration—the process of combining criterion scores into overall impact scores—disadvantage black applicants. Using multilevel models and matching on key variables including career stage, gender, and area of science, we find little evidence for racial disparities emerging in the process of combining preliminary criterion scores into preliminary overall impact scores. Instead, preliminary criterion scores fully account for racial disparities—yet do not explain all of the variability—in preliminary overall impact scores.

Re. invited talks:

Ford, H. L., Brick, C., Blaufuss, K., & Dekens, P. S. (2018). Gender inequity in speaking opportunities at the American Geophysical Union Fall Meeting. *Nature Communications*, 9(1), 1-6.

Abstract. Implicit and explicit biases impede the participation of women in science, technology, engineering, and mathematic (STEM) fields. Across career stages, attending conferences and presenting research are ways to spread scientific results, find job opportunities, and gain awards. Here, we present an analysis by gender of the American Geophysical Union Fall Meeting speaking opportunities from 2014 to 2016. We find that women were invited and assigned oral presentations less often than men. However, when we control for career stage, we see similar rates between women and men and women sometimes outperform men. At the same time, women elect for poster presentations more than men. Male primary conveners allocate invited abstracts and oral presentations to women less often and below the proportion of women authors. These results highlight the need to provide equal opportunity to women in speaking roles at scientific conferences as part of the overall effort to advance women in STEM.

Re. professional achievements:

Vaid, J., & Geraci, L. (2016). An examination of women's professional visibility in cognitive psychology. *Feminism & Psychology*, 26(3), 292-319.

Abstract. Mainstream psychological research has been characterized as androcentric in its construction of males as the norm. Does an androcentric bias also characterize the professional visibility of psychologists? We examined this issue for cognitive psychology, where the gender distribution in doctoral degrees has been roughly equal for several decades. Our investigation revealed that, across all indicators surveyed, male cognitive psychologists are more visible than their female counterparts: they are over-represented in professional society governance, as editors-in-chief of leading journals in the field, as Fellows in professional societies, and as recipients of prestigious senior level awards. Taken together, our findings indicate that a gender parity in doctoral degrees in cognitive psychology does not translate into a parity in professional visibility. We discuss a variety of potential reasons for the observed gender gap and suggest that, without attention to gendered structures of status and power, as noted by Shields, existing gender hierarchies may persist and be reproduced.

Recognize that where a scholar lands is more important than where they trained.

Way, S. F., Morgan, A. C., Larremore, D. B., & Clauset, A. (2019). Productivity, prominence, and the effects of academic environment. *Proceedings of the National Academy of Sciences* 116(22), 10729-10733.

Abstract. Faculty at prestigious institutions produce more scientific papers, receive more citations and scholarly awards, and are typically trained at more-prestigious institutions than faculty with less prestigious appointments. This imbalance is often attributed to a meritocratic system that sorts individuals into more-prestigious positions according to their reputation, past achievements, and potential for future scholarly impact. Here, we investigate the determinants of scholarly productivity and measure their dependence on past training and current work environments. To distinguish the effects of these environments, we apply a matched-pairs experimental design to career and productivity trajectories of 2,453 early-career faculty at all 205 PhD-granting computer science departments in the United States and Canada, who together account for over 200,000 publications and 7.4 million citations. Our results show that the prestige of faculty's current work environment, not their training environment, drives their future scientific productivity, while current and past locations drive prominence. Furthermore, the characteristics of a work environment are more predictive of faculty productivity and impact than mechanisms representing preferential selection or retention of more-productive scholars by more-prestigious departments. These results identify an environmental mechanism for cumulative advantage, in which an individual's past successes are "locked in" via placement into a more prestigious environment, which directly facilitates future success. The scientific productivity of early-career faculty is thus driven by where they work, rather than where they trained for their doctorate, indicating a limited role for doctoral prestige in predicting scientific contributions.

Beware of bias in recommendation letters.

Re. excellence:

Dutt, K., Pfaff, D. L., Bernstein, A. F., Dillard, J. S., & Block, C. J. (2016). Gender differences in recommendation letters for postdoctoral fellowships in geoscience. *Nature Geoscience*, 9(11), 805-808.

Abstract. Gender disparities in the fields of science, technology, engineering and mathematics, including the geosciences, are well documented and widely discussed^{1,2}. In the geosciences, despite receiving 40% of doctoral degrees, women hold less than 10% of full professorial positions³. A significant leak in the pipeline occurs during postdoctoral years⁴, so biases embedded in postdoctoral processes, such as biases in recommendation letters, may be deterrents to careers in geoscience for women. Here we present an analysis of an international data set of 1,224 recommendation letters, submitted by recommenders from 54 countries, for postdoctoral fellowships in the geosciences over the period 2007–2012. We examine the relationship between applicant gender and two outcomes of interest: letter length and letter tone. Our results reveal that female applicants are only half as likely to receive excellent letters versus good letters compared to male applicants. We also find no evidence that male and female recommenders differ in their likelihood to write stronger letters for male applicants over female applicants. Our analysis also reveals significant regional differences in letter length, with letters from the Americas being significantly longer than any other region, whereas letter tone appears to be distributed equivalently across all world regions. These results suggest that women are significantly less likely to receive excellent recommendation letters than their male counterparts at a critical juncture in their career.

Re. doubt raising:

Madera, J. M., Hebl, M. R., Dial, H., Martin, R., & Valian, V. (2019). Raising doubt in letters of recommendation for academia: Gender differences and their impact. *Journal of Business and Psychology*, 34(3), 287-303.

Abstract. The extent of gender bias in academia continues to be an object of inquiry, and recent research has begun to examine the particular gender biases emblematic in letters of recommendations. This current two-part study examines differences in the number of doubt raisers that are written in 624 authentic letters of recommendations for 174 men and women applying for eight assistant professor positions (study 1) and the impact of these doubt raisers on 305 university professors who provided evaluations of recommendation letters (study 2). The results show that both male and female recommenders use more doubt raisers in letters of recommendations for women compared to men and that the presence of certain types of doubt raisers in letters of recommendations results in negative outcomes for both genders. Since doubt raisers are more frequent in letters for women than men, women are at a disadvantage relative

to men in their applications for academic positions. We discuss the implications and need for additional future research and practice that (1) raises awareness that letter writers are gatekeepers who can improve or hinder women's progress and (2) develops methods to eliminate the skewed use of doubt raisers.

Beware of bias in teaching evaluations.

Boring, Anne, Kelli Ottoboni, and Philip B. Stark (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. ScienceOpen Research.

Abstract. Student evaluations of teaching (SET) are widely used in academic personnel decisions as a measure of teaching effectiveness. We show:

- SET are biased against female instructors by an amount that is large and statistically significant.
- The bias affects how students rate even putatively objective aspects of teaching, such as how promptly assignments are graded.
- The bias varies by discipline and by student gender, among other things.
- It is not possible to adjust for the bias, because it depends on so many factors.
- SET are more sensitive to students' gender bias and grade expectations than they are to teaching effectiveness.
- Gender biases can be large enough to cause more effective instructors to get lower SET than less effective instructors.

These findings are based on nonparametric statistical tests applied to two datasets: 23,001 SET of 379 instructors by 4,423 students in six mandatory first-year courses in a five-year natural experiment at a French university, and 43 SET for four sections of an online course in a randomized, controlled, blind experiment at a US university.

Conclusion. In two very different universities and in a broad range of course topics, SET measure students' gender biases better than they measure the instructor's teaching effectiveness. Overall, SET disadvantage female instructors. There is no evidence that this is the exception rather than the rule. Hence, the onus should be on universities that rely on SET for employment decisions to provide convincing affirmative evidence that such reliance does not have disparate impact on women, underrepresented minorities, or other protected groups. Because the bias varies by course and institution, affirmative evidence needs to be specific to a given course in a given department in a given university. Absent such specific evidence, SET should not be used for personnel decisions.

Know the value of all types of service.

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.

Abstract. This paper investigates the amount of academic service performed by female versus male faculty. We use 2014 data from a large national survey of faculty at more than 140 institutions as well as 2012 data from an online annual performance reporting system for tenured and tenure-track faculty at two campuses of a large public, Midwestern University. We find evidence in both data sources that, on average, women faculty perform significantly more service than men, controlling for rank, race/ethnicity, and field or department. Our analyses suggest that the male-female differential is driven more by internal service—i.e., service to the university, campus, or department—than external service—i.e., service to the local, national, and international communities—although significant heterogeneity exists across field and discipline in the way gender differentials play out.

Know the different ways diversity work manifests.

Re. diversity statements:

Sylvester, C. Y. C., Sánchez-Parkinson, L., Yettaw, M., & Chavous, T. (2019). The promise of diversity statements: Insights and a framework developed from faculty applications. *The National Center for Institutional Diversity (NCID) Currents*, 1(1).

Introduction. As students in colleges and universities continue to diversify along myriad dimensions, there is a need to hire faculty who have the expertise, knowledge, and commitments needed to foster intellectually and culturally rich, inclusive, and equitable learning environments. Faculty can make these contributions in multiple ways. ... Here we present a case example that reflects the above principles, for which we analyzed diversity statements written by faculty applicants across social sciences, humanities, and natural sciences departments in a college of liberal arts within a large research university. Our comprehensive analysis of the content and features of applicants' diversity statements informed the development of an initial diversity statement framework, providing an organization for the numerous ways an applicant might articulate their demonstrated contributions to DEI. This type of framework can be useful to faculty applicants as they approach writing their DEI statements and for institutions that request DEI statements to support their ability to effectively identify and evaluate information about desirable skills, knowledge, and perspectives for their faculty roles.

Re. considering societal impact:

Zárate, M. A., Hall, G. N., & Plaut, V. C. (2017). Researchers of color, fame, and impact. *Perspectives on Psychological Science*, 12(6), 1176-1178.

Abstract. Fame and eminence, as traditionally measured, limit the definition of impact to the publication world. We add two types of impact to the traditional measures of fame and eminence. Many of the traditional measures of fame or eminence are based on social-network connections, whereby individuals appoint other people to positions of eminence. Editorial boards are one specific example. Eminence is also limited to number of publications, for example, with little regard for the impact of those publications at the societal level. In addition to the dominant measures of eminence, societal impact broadens the definition of impact to reflect real-world changes. Two examples include mentoring, which is rarely mentioned as a criterion for eminence, and policy value, such as when research influences important public policy. These additions are discussed in reference to the general underrepresentation of researchers of color in academia.

Ensure fairness during job talks.

Blair-Loy, Mary et al. 2017. Gender in Engineering Departments: Are There Gender Differences in Interruptions of Academic Job Talks? *Social Sciences*. 6(29).

Abstract. We use a case study of job talks in five engineering departments to analyze the under-studied area of gendered barriers to finalists for faculty positions. We focus on one segment of the interview day of short-listed candidates invited to campus: the “job talk”, when candidates present their original research to the academic department. We analyze video recordings of 119 job talks across five engineering departments at two Research 1 universities. Specifically, we analyze whether there are differences by gender or by years of post-Ph.D. experience in the number of interruptions, follow-up questions, and total questions that job candidates receive. We find that, compared to men, women receive more follow-up questions and more total questions. Moreover, a higher proportion of women’s talk time is taken up by the audience asking questions. Further, the number of questions is correlated with the job candidate’s statements and actions that reveal he or she is rushing to present their slides and complete the talk. We argue that women candidates face more interruptions and often have less time to bring their talk to a compelling conclusion, which is connected to the phenomenon of “stricter standards” of competence demanded by evaluators of short-listed women applying for a masculine-typed job. We conclude with policy recommendations.