



Abstracts on Bias in Faculty Evaluation

Compiled by the UMD ADVANCE Program (2022)

Metrics and Indicators of Scholarly Productivity.....	2
Epistemic Exclusion and Narrow Conceptions of Scholarship	4
Rigidity in Timelines.....	6
Assigning Credit in Group Work and Collaborations	7
Teaching Evaluations	8
Counting Diversity, Equity, and Inclusion Work	10



Metrics and Indicators of Scholarly Productivity

Mitchneck, B. (2020). *Synthesizing Research on Gender Biases and Intersectionality in Citation Analysis and Practices*. ARC Network. <https://www.equityinstem.org/virtual-visiting-scholars/mitchneck>

This meta synthesis of the multidisciplinary literature on citation practice and citation indices includes 100 different sources including peer-reviewed publications, books, websites and blogs. Although the meta synthesis was intersectional in framing, there is relatively little direct engagement within the literature about women of color and citation practice. The meta synthesis assesses 1) literature about influences over citation practice and citation indices; and 2) literature on those factors and how they impact the likelihood that faculty of color are on the receiving end of citation practice. The meta synthesis identifies key areas of thought including underlying theoretical frameworks to understand citation practice, the social influences over citation practice, and alternative ways to measure research impact and productivity. The practice of citing and the resulting citation indices generally reflect patterns of gender and racial biases found in the literature on evaluation and the underrepresentation of women in science. Recent publications highlight the ways that the lack of visibility and efforts to create an invisibility of men and women of color in STEM may be overcome by using big data methods and technologies. These promising methods can and should be used to analyze directly the intersecting roles of gender, race, and ethnicity of the person being cited and the person doing the citing into summary measures of productivity and impact. At a minimum, this meta synthesis finds that citation practice and thus citation indices are not normative measures of scholarly productivity and impact but are highly influenced by any number of factors in addition to merit and quality and are subject to any number of ways that they are consciously and unconsciously manipulated to the disadvantage of out groups. Thus, the use of citation indices as single measures of quality and impact directly embed biases into our standard measure of merit.

Geraci, L., Balsis, S., & Busch, A. J. B. (2015). *Gender and the h index in psychology*. *Scientometrics*, *105*(3), 2023-2034.

It has become increasingly common to rely on the *h* index to assess scientists' contributions to their fields, and this is true in psychology. This metric is now used in many psychology departments and universities to make important decisions about hiring, promotions, raises, and awards. Yet, a growing body of research shows that there are gender differences in citations and *h* indices. We sought to draw attention to this literature, particularly in psychology. We describe the presence of a gender effect in *h* index in psychology and analyze why the effect is important to consider. To illustrate the importance of this effect, we translate the observed gender effect into a meaningful metric—that of salary—and show that the gender difference in *h* index could translate into significant financial costs for female faculty. A variety of factors are discussed that have been shown to give rise to gender differences in impact. We conclude that the *h* index, like many other metrics, may reflect systematic gender differences in academia, and we suggest using caution when relying on this metric to promote and reward academic psychologists.

West, R. E., & Rich, P. J. (2012). *Rigor, impact and prestige: A proposed framework for evaluating scholarly publications*. *Innovative Higher Education*, *37*(5), 359-371.

As publication pressure has increased in the world of higher education, more journals, books, and other publication outlets have emerged. Thus it is critical to develop clear criteria for effectively evaluating the quality of publication outlets. Without such criteria funding agencies and promotion committees are



forced to guess at how to evaluate a scholar's portfolio. In this article, we explore the perils of evaluating journals based on a single quantitative measure (e.g., the Impact Factor rating of the Institute for Science Information). We then discuss key considerations for evaluating scholarship, including three main criteria: rigor, impact, and prestige. Finally, we conclude with examples of how these criteria could be applied in evaluating scholarship.

Hofstra, B., Kulkarni, V.V., Galvez, S., He, B., Jurafsky, D. & McFarland, D. (2020). The diversity-innovation paradox in science. PNAS, 117(17), 9284-9291.

By analyzing data from nearly all US PhD recipients and their dissertations across three decades, this paper finds demographically underrepresented students innovate at higher rates than majority students, but their novel contributions are discounted and less likely to earn them academic positions. The discounting of minorities' innovations may partly explain their underrepresentation in influential positions of academia.



Epistemic Exclusion and Narrow Conceptions of Scholarship

Settles, I. H., Jones, M. K., Buchanan, N. T., & Dotson, K. (2020). Epistemic exclusion: Scholar(ly) devaluation that marginalizes faculty of color. *Journal of Diversity in Higher Education, 14*(4), 493-507.

Faculty of color experience a number of challenges within academia, including tokenism, marginalization, racial microaggressions, and a disconnect between their racial/ethnic culture and the culture within academia. The present study examined epistemic exclusion as another challenge in which formal institutional systems of evaluation combine with individual biases toward faculty of color to devalue their scholarship and deem them illegitimate as scholars. Using data from interviews with 118 faculty of color from a single predominantly White, research-intensive institution, we found that epistemic exclusion occurs through formal hierarchies that determine how scholarship is valued and the metrics used to assess quality, and through informal processes that further convey to faculty of color that they and their scholarship are devalued. In addition, there was variability in reporting these experiences by race, gender, nationality, and discipline. We found that faculty of color coped with epistemic exclusion by being assertive and by seeking validation and support outside the institution. Finally, participants described a number of negative work-related and psychological consequences of their epistemic exclusion. We discuss epistemic exclusion as a form of academic gatekeeping that impedes the recruitment, advancement, and retention of faculty of color and offer strategies to address this barrier.

Gonzales, L. D. (2018). Subverting and minding boundaries: The intellectual work of women. *The Journal of Higher Education, 89*(5), 677-701.

Using various methods and analytical angles, researchers consistently show that members of non-dominant groups, including women, experience academia as a hostile and marginalizing space. Such work is important, and yet, it is equally important that researchers approach the study of women's academic careers by elevating their intellectual labor. In this study, I take up two questions: (1) What are the origins of women's intellectual work and (2) How do women go about doing their intellectual work? My findings suggest that women tend to locate the origins of their work in the everyday rather than in formal educational sites, such as disciplinary contexts or classrooms. In terms of the doing of their intellectual work, I found that most women utilize subversive tactics, as they challenge disciplinary and professional boundaries that have historically governed the recognition and legitimation of knowledge within academe. However, drawing from critical race feminism, I also find some notable distinctions between Women of Color and White women, and suggest that future researchers attend more carefully to how power and privilege yields particular conditions and consequences among women. This paper offers important insights for peer reviewers (e.g., hiring, promotion, disciplinary award committees, and publication reviewers) as to the grounding(s) and distinctive contribution(s) of women's intellectual work.

O'Meara, K., Eatman, T., & Petersen, S. (2015). Advancing engaged scholarship in promotion and tenure: A roadmap and call for reform. *Liberal Education, 101*(3), n3.

Despite the precipitous increase in nontenure-track faculty appointments, the promotion and tenure process continues to operate as a central "motivational and cultural force in the academic lives" of many faculty members. As a part of larger reward systems, the promotion and tenure process reflects institutional values, aspirations, privileges, and power structures. In 2010, the authors began collaborating



ADVANCE

Investing in Faculty Success and Inclusive Work Environments

with campus teams interested in reforming their institutions' promotion and tenure guidelines in order to define, assess, document, and reward engaged scholarship. The organizing vehicle since 2010 has been the Faculty Rewards Institute at the annual Eastern Region Campus Compact conference. To date the authors have hosted 41 campus teams, including 116 individual participants. They designed the daylong institute as an opportunity to share knowledge and tools and to enhance collective critical agency around campus reform related to faculty roles and rewards, with particular attention to how current promotion and tenure policies may exclude engaged scholarship and engaged faculty. In this article, the authors begin by sharing the process they use to engage campus teams and individuals in diagnosing what is wrong within the promotion and tenure system, what they want to change, and how to make that change. They share a template for studying and reforming promotion and tenure policies to ensure that they better reward engaged scholarship.



Rigidity in Timelines

Cech, E. A., & Blair-Loy, M. (2014). Consequences of flexibility stigma among academic scientists and engineers. *Work and Occupations*, 41(1), 86-110.

Flexibility stigma, the devaluation of workers who seek or are presumed to need flexible work arrangements, fosters a mismatch between workplace demands and the needs of professionals. The authors survey “ideal workers”—science, technology, engineering, and math faculty at a top research university—to determine the consequences of working in an environment with flexibility stigma. Those who report this stigma have lower intentions to persist, worse work–life balance, and lower job satisfaction. These consequences are net of gender and parenthood, suggesting that flexibility stigma fosters a problematic environment for many faculty, even those not personally at risk of stigmatization.

Damaske, S., Ecklund, E. H., Lincoln, A. E., & White, V. J. (2014). Male scientists’ competing devotions to work and family: Changing norms in a male-dominated profession. *Work and Occupations*, 41(4), 477-507.

Using in-depth interviews with 74 men across different ranks in biology and physics at prestigious U.S. universities, the authors ask to what extent changing norms of fatherhood and a flexible workplace affect men working in a highly male-dominated profession and what variation exists in family forms. The authors conceptualize four typologies of men: those forgoing children, egalitarian partners, neotraditional dual earners, and traditional breadwinners. Findings suggest male scientists hold strong work devotions, yet a growing number seek egalitarian relationships, which they frame as reducing their devotion to work. The majority of men find the all-consuming nature of academic science conflicts with changing fatherhood norms.

Sallee, M. W. (2012). The ideal worker or the ideal father: Organizational structures and culture in the gendered university. *Research in Higher Education*, 53(7), 782- 802.

While literature has focused on the ways in which organizational structures exclude women from the workplace, this article suggests that the inverse is also true: organizational structures and culture prevent men from being involved in the home. Using theories of gendered organizations as a guide, this article draws on interviews with 70 faculty fathers at four research universities to explore the tension that many men feel navigating their responsibilities in the home while simultaneously aiming to fulfill the norms of the ideal worker, which holds that employees are always available to perform work and have few responsibilities in the home. Data suggest that institutions and those within them penalize men who appear too committed to their families. Some participants crafted identities for themselves that separated their roles as professor and father while others struggled to reconcile their two roles. In short, institutional structures and culture play a critical role in shaping faculty identity, both on and off-campus.



Assigning Credit in Group Work and Collaborations

Sarsons, H., Gërkhani, K., Reuben, E., & Schram, A. (2021). Gender differences in recognition for group work. *Journal of Political Economy*, 129(1), 101-147.

We study whether gender influences credit attribution for group work using observational data and two experiments. We use data from academic economists to test whether coauthorship matters differently for tenure for men and women. We find that, conditional on quality and other observables, men are tenured similarly regardless of whether they coauthor or solo author. Women, however, are less likely to receive tenure the more they coauthor. We then conduct two experiments that demonstrate that biases in credit attribution in settings without confounds exist. Taken together, our results are best explained by gender and stereotypes influencing credit attribution for group work.

Haynes, M. C., & Heilman, M. E. (2013). It had to be you (not me)! Women’s attributional rationalization of their contribution to successful joint work outcomes. *Personality and Social Psychology Bulletin*, 39(7), 956-969.

We investigated the tendency of women to undervalue their contributions in collaborative contexts. Participants, who believed they were working with another study participant on a male sex-typed task, received positive feedback about the team’s performance. Results indicated that women and men allocated credit for the joint success very differently. Women gave more credit to their male teammates and took less credit themselves unless their role in bringing about the performance outcome was irrefutably clear (Studies 1 and 2) or they were given explicit information about their likely task competence (Study 4). However, women did not credit themselves less when their teammate was female (Study 3). Together these studies demonstrate that women devalue their contributions to collaborative work, and that they do so by engaging in *attributional rationalization*, a process sparked by women’s negative performance expectations and facilitated by source ambiguity and a satisfactory “other” to whom to allocate credit.

Knobloch-Westerwick, S., Glynn, C. J., & Huge, M. (2013). The Matilda effect in science communication: an experiment on gender bias in publication quality perceptions and collaboration interest. *Science communication*, 35(5), 603-625.

An experiment with 243 young communication scholars tested hypotheses derived from role congruity theory regarding impacts of author gender and gender typing of research topics on perceived quality of scientific publications and collaboration interest. Participants rated conference abstracts ostensibly authored by females or males, with author associations rotated. The abstracts fell into research areas perceived as gender-typed or gender-neutral to ascertain impacts from gender typing of topics. Publications from male authors were associated with greater scientific quality, in particular if the topic was male-typed. Collaboration interest was highest for male authors working on male-typed topics. Respondent sex did not influence these patterns.



Teaching Evaluations

Kreitzer, R. J., & Sweet-Cushman, J. (2021). Evaluating student evaluations of teaching: a review of measurement and equity bias in SETs and recommendations for ethical reform. *Journal of Academic Ethics*, 1-12.

Student evaluations of teaching are ubiquitous in the academe as a metric for assessing teaching and frequently used in critical personnel decisions. Yet, there is ample evidence documenting both measurement and equity bias in these assessments. Student Evaluations of Teaching (SETs) have low or no correlation with learning. Furthermore, scholars using different data and different methodologies routinely find that women faculty, faculty of color, and other marginalized groups are subject to a disadvantage in SETs. Extant research on bias on teaching evaluations tend to review only the aspect of the literature most pertinent to that study. In this paper, we review a novel dataset of over 100 articles on bias in student evaluations of teaching and provide a nuanced review of this broad but established literature. We find that women and other marginalized groups do face significant biases in standard evaluations of teaching – however, the effect of gender is conditional upon other factors. We conclude with recommendations for the judicious use of SETs and avenues for future research.

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

Many universities use student evaluations of teachers (SETs) as part of consideration for tenure, compensation, and other employment decisions. However, in doing so, they may be engaging in discriminatory practices against female academics. This study further explores the relationship between gender and SETs described by MacNeill, Driscoll, and Hunt (2015) by using both content analysis in student-evaluation comments and quantitative analysis of students' ordinal scoring of their instructors. The authors show that the language students use in evaluations regarding male professors is significantly different than language used in evaluating female professors. They also show that a male instructor administering an identical online course as a female instructor receives higher ordinal scores in teaching evaluations, even when questions are not instructor-specific. Findings suggest that the relationship between gender and teaching evaluations may indicate that the use of evaluations in employment decisions is discriminatory against women.

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.

Student evaluations of teaching are widely believed to contain gender bias. In this study, we conduct a randomized experiment with the student evaluations of teaching in four classes with large enrollments, two taught by male instructors and two taught by female instructors. In each of the courses, students were randomly assigned to either receive the standard evaluation instrument or the same instrument with language intended to reduce gender bias. Students in the anti-bias language condition had significantly higher rankings of female instructors than students in the standard treatment. There were no differences between treatment groups for male instructors. These results indicate that a relatively simple intervention in language can potentially mitigate gender bias in student evaluation of teaching.



Storage, D., Horne, Z., Cimpian, A., & Leslie, S. J. (2016). The frequency of “Brilliant” and “Genius” in teaching evaluations predicts the representation of women and African Americans across fields. *PloS one*, 11(3), e0150194.

Women and African Americans—groups targeted by negative stereotypes about their intellectual abilities—may be underrepresented in careers that prize brilliance and genius. A recent nationwide survey of academics provided initial support for this possibility. Fields whose practitioners believed that natural talent is crucial for success had fewer female and African American PhDs. The present study seeks to replicate this initial finding with a different, and arguably more naturalistic, measure of the extent to which brilliance and genius are prized within a field. Specifically, we measured field-by-field variability in the emphasis on these intellectual qualities by tallying—with the use of a recently released online tool—the frequency of the words “brilliant” and “genius” in over 14 million reviews on RateMyProfessors.com, a popular website where students can write anonymous evaluations of their instructors. This simple word count predicted both women’s and African Americans’ representation across the academic spectrum. That is, we found that fields in which the words “brilliant” and “genius” were used more frequently on RateMyProfessors.com also had fewer female and African American PhDs. Looking at an earlier stage in students’ educational careers, we found that brilliance-focused fields also had fewer women and African Americans obtaining bachelor’s degrees. These relationships held even when accounting for field-specific averages on standardized mathematics assessments, as well as several competing hypotheses concerning group differences in representation. The fact that this naturalistic measure of a field’s focus on brilliance predicted the magnitude of its gender and race gaps speaks to the tight link between ability beliefs and diversity.



Counting Diversity, Equity, and Inclusion Work

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.

Little work has addressed the ways in which race and gender intersect and shape Black professors' experiences as they seek professional advancement. Framed by critical race theory, this qualitative study uses discourse analysis to analyze the narratives of 28 Black professors employed at two research universities. Findings suggest that faculty perceive race and gender influencing their evaluations for academic advancement, with key gender distinctions in discourses about teaching and service in relation to professional success. Black women appear to experience demands in these domains as more emotionally and physically taxing than their male counterparts, adding strain to the tenure and advancement process.

Jimenez, M. F., Laverty, T. M., Bombaci, S. P., Wilkins, K., Bennett, D. E., & Pejchar, L. (2019). Underrepresented faculty play a disproportionate role in advancing diversity and inclusion. *Nature Ecology & Evolution*, 3(7), 1030-1033.

A diverse and inclusive scientific community is more productive, innovative and impactful, yet ecology and evolutionary biology continues to be dominated by white male faculty. We quantify faculty engagement in activities related to diversity and inclusion and identify factors that either facilitate or hinder participation. Through a nationwide survey, we show that faculty with underrepresented identities disproportionately engage in diversity and inclusion activities, yet such engagement was not considered important for tenure. Faculty perceived time and funding as major limitations, which suggests that institutions should reallocate resources and reconsider how faculty are evaluated to promote shared responsibility in advancing diversity and inclusion.

Misra, J., Kuvaeva, A., O'Meara, K., Culpepper, D. K., & Jaeger, A. (2021). Gendered and racialized perceptions of faculty workloads. *Gender & Society*, 35(3), 358-394.

Faculty workload inequities have important consequences for faculty diversity and inclusion. On average, women faculty spend more time engaging in service, teaching, and mentoring, while men, on average, spend more time on research, with women of color facing particularly high workload burdens. We explore how faculty members perceive workload in their departments, identifying mechanisms that can help shape their perceptions of greater equity and fairness. White women perceive that their departments have less equitable workloads and are less committed to workload equity than white men. Women of color perceive that their departments are less likely to credit their important work through departmental rewards systems than white men. Workload transparency and clarity, and consistent approaches to assigning classes, advising, and service, can reduce women's perceptions of inequitable and unfair workloads. Our research suggests that departments can identify and put in place a number of key practices around workload that will improve gendered and racialized perceptions of workload.