

Gender Bias in Peer Review of Faculty

Notes for DACOPAT by Simon Peacock and Anne Condon, February 2009

We value the principle that merit, as revealed by the process of peer review, drives our tenure and promotion decisions. However, it is well documented that unconscious bias or assumptions regarding gender can negatively colour evaluation of scientific accomplishment [1-5]. Following are some points (adapted from Ref. 1) to keep in mind as we evaluate tenure and promotion cases and particularly letters of evaluation.

First, studies have shown that **letters of recommendation for women differ systematically from those for men**. For example, Trix and Psenka [2] examined over 300 letters of recommendation for medical faculty at a large American medical school in the mid-90's. Their analysis showed that, compared with letters for men, the letters for women were more likely to:

1. *be shorter*, yet contain more references to aspects of personal life;
2. *provide minimal assurance*, with comments on generalities but without specifics (15% of letters for females vs 6% for males);
3. *be doubt raisers*, by including hedges, potentially negative comments, unexplained comments, faint praise, or irrelevancies (24% vs 12%);
4. *use stereotypic attributes*, such as “compassionate” or “relates well” rather than “successful” or nouns such as “accomplishment” or “achievement”; and
5. *use grindstone adjectives*, such as “hardworking” or “conscientious” (34% vs 23%).

Additionally, letters for women had *less repetition of standout adjectives*, such as “outstanding”.

Second, **evaluations of academic CV's can be influenced by gender** [3]. Academics in Psychology in the U.S. were asked to evaluate a CV of a job applicant, with CV's being identical except that each was randomly assigned either a male or female name. Both men and women who evaluated the CV were significantly more likely to recommend hiring the “male” applicant than the “female” applicant.

Third, **perceptions of scientific productivity can be influenced by gender**. Wennerås and Wold [4] examined the peer-review scores of applicants for postdoctoral fellowships from the Swedish Medical Research Council in 1995. They correlated reviewers' “competence” scores of perceived scientific productivity with several measures of scientific impact, including total number of publications, total number of first-author publications, sum of impact factors of journal articles (“total impact”), sum of impact factors of first-author journal articles, sum of citations, and sum of first-author citations. They found that reviewers assigned lower competence scores to women than to men with the same measure of scientific impact. A multiple-regression analysis revealed that gender exerted more influence on assigned competence scores than other factors, such as scientific field or university affiliation.

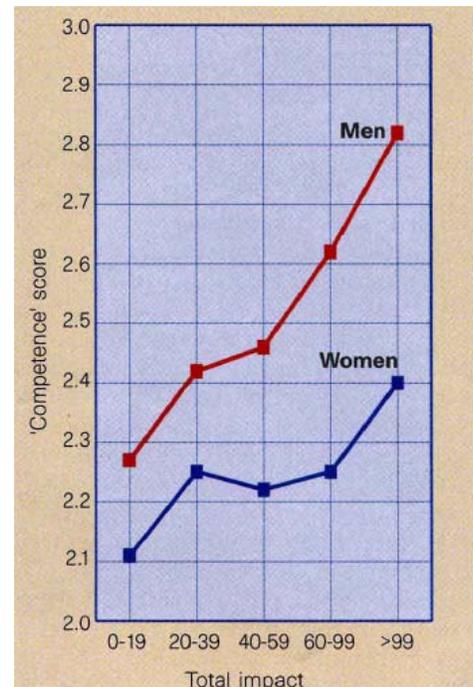


Figure 1 of Wennerås and Wold [4], showing mean competence score as a function of total impact, for men and women.

References

1. Reviewing Applicants: Research on Bias and Assumptions. Women in Science and Engineering Leadership Institute, U. Wisconsin-Madison, 2006.
2. Exploring the Color of Glass: Letters of Recommendation for Female and Male Medical Faculty. F. Trix and C. Psenka, *Discourse and Society*, 2003.
3. The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study. Steinpreis et al., *Sex Roles*, 1999.
4. Nepotism and Sexism in Peer Review. C. Wennerås and A. Wold, *Nature* 387, 1997
5. Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering. A report of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine of the National Academies, the National Academies Press, Washington, D.C., 2007.

Unconscious Bias in Faculty and Leadership Recruitment: A Literature Review

Although women and minorities have made significant strides in achieving equality in the workplace, they are still underrepresented in the upper strata of organizations, including senior faculty and leadership positions at medical schools and teaching hospitals.^{a,1,2} Within the last decade, social science researchers have pursued the theory of “unconscious bias” as one barrier to workplace equality that may persist despite a general commitment to increase diversity across the academic medicine workforce and other organizations. This *Analysis in Brief* reviews the scientific literature on the theory of unconscious bias, explores the role of unconscious bias in job recruitment and evaluations, and offers suggestions for search committees and others involved in hiring decisions at medical schools and teaching hospitals.

Background and Method

Unconscious bias refers to social stereotypes about certain demographics or groups of people that individuals form outside of their own conscious awareness. Social scientists argue that most people have some degree of unconscious bias because it stems from our natural tendency to make associations to help us organize our social worlds.³

To analyze and present the most relevant studies on unconscious bias, I performed an extensive literature search and sorted articles by their research methodologies and focus.^b The eight articles selected for this *Analysis in Brief* were specifically about career-related unconscious bias, except for those on the Implicit

Table 1. Experimental Studies of Career-related Unconscious Bias*

Method	Results
Steinpres, Anders, & Ritzke (1999)	
Participants (238 academic psychologists) rated the <i>curriculum vitae</i> (CVs) of real-life psychologists: one CV from early career and one CV from later career. They received one of these CVs with either a male or female name on it, and were asked whether they would hire the job applicant (for those who received the early-career CV) or whether they would award the candidate tenure (for those who received the later-career CV).	<ul style="list-style-type: none"> • Both male and female participants were more likely to hire the male applicant than the female applicant and were more likely to report that the male applicant had adequate experience in research, teaching, and service. • Participants were four times more likely to write comments of concern in the margins of their questionnaires for female tenure candidates than for male tenure candidates.
Heilman & Okimoto (2007)	
Given descriptions of fictitious male and female managers who were successful in male-dominated jobs, undergraduate students evaluated the managers on several measures, including likeability, interpersonal hostility, competence, and desirability as a boss.	Female managers were rated as less likeable, more interpersonally hostile, less competent, and less desirable as bosses than were male managers.
King, Madera, Hebl, & Knight (2006)	
Participants (155 white males) viewed a fictitious resume of an Asian-American, Mexican-American, African-American, or Caucasian man that was either high or low quality. They gave the target of the resume an overall evaluation (based on intelligence, motivation, and likelihood to be hired) and a rating on suitability for various occupations (categorized as high- or low-status jobs).	<ul style="list-style-type: none"> • On the overall evaluation, African-American targets were rated the least positively and Asian-American targets were rated the most positively. • Asian-American targets were rated as the most suitable for high-status jobs, whereas Mexican-American targets were rated as the least suitable for high-status jobs. Mexican-American targets were rated as the most suitable for low-status jobs, whereas Caucasian targets were rated as the least suitable for low-status jobs.
Biernat & Manis (1994)	
Participants (143 white undergraduates) viewed 40 photographs of black and white individuals, each paired with definitions of two words. The participants rated the verbal ability of the photographed individuals as if those individuals had provided the definitions.	Black individuals were rated as having lower verbal ability than were white individuals, suggesting that the participants had a bias that black people are less verbally skilled than white people.

* For complete bibliographical information on the literature in this table, see www.aamc.org/data/aib.

Association Test (IAT), which were included as background material because of their prevalence in the literature and their direct illustration of unconscious bias. Articles on career-related unconscious bias included methodologies of two types: experimental studies involving artificial workplace situations and examinations of unconscious bias in actual hiring and review practices. The implications about recruitment practices in academic medicine drawn from this literature review are gener-

alized from the findings of unconscious bias in many fields.

Results and Discussion

Implicit Association Test

Though not directly connected to career-related unconscious bias, studies involving the IAT receive the most support for the theory of unconscious bias in general. The IAT has consistently demonstrated that people unconsciously prefer white over black, young over old, and thin over fat, and that people have stereotypic associations

^a For the numbered list of references, see www.aamc.org/data/aib.

^b For more information on the methodology of the literature search, see www.aamc.org/data/aib.

Table 2. Real-life Examinations of Career-related Unconscious Bias*

Method	Results
Trix & Psenka (2003)	
The researchers analyzed 312 real letters of recommendation that helped medical school faculty receive their clinical and research positions. The letters were received by a large U.S. medical school from 1992 to 1995.	Compared with letters of recommendation for males, letters for females were shorter, were more likely to be "letters of minimal assurance" (e.g., lacking in specificity), were more likely to contain gender terms (e.g., "she is an intelligent young lady"), and were more likely to include "doubt raisers" (e.g., criticisms, hedges, faint praise).
Wennerås & Wold (1997)	
The researchers evaluated whether the peer-review system of postdoctoral fellowships at the Swedish Medical Research Council was biased against women.	For peer-reviewers to equally rate men and women on scientific competency, women needed an equivalent of approximately three more articles in <i>Nature</i> or <i>Science</i> , or 20 more articles in a specialist journal, such as <i>Neuroscience</i> or <i>Radiology</i> .
Goldin & Rouse (2000)	
To test for sex-biased hiring in symphony orchestras, the researchers compared two audition procedures: "blind" auditions (adopted in 1970—involves the use of screens to conceal candidates' identities) and "not-blind" auditions (no use of screens).	<ul style="list-style-type: none"> Blind auditions increases the likelihood that a female will be hired by 25 percent. The switch to blind auditions in 1970 explains 30 percent of the increase in the proportion of females among new hires.
Bertrand & Mullainathan (2003)	
To examine the effect of race on receiving job callbacks, the researchers responded with fictitious resumes to help-wanted ads in Boston and Chicago newspapers. The resumes were altered from actual ones found on job search Web sites. The researchers categorized the new resumes as high or low quality and assigned them an equal number of traditionally black names (e.g., Lakisha) or traditionally white names (e.g., Greg).	<ul style="list-style-type: none"> Resumes with white names had a 50 percent greater chance of receiving a callback than did resumes with black names (10.08% vs. 6.70%, respectively). Higher-quality resumes elicited 30 percent more callbacks for whites, whereas they only elicited 9 percent more callbacks for blacks. Employers who listed "Equal Opportunity Employer" in their ad discriminated just as much as other employers.

* For complete bibliographical information on the literature in this table, see www.aamc.org/data/aib.

linking males with science and careers and females with liberal arts and family.^{4,5} Typically, the IAT requires test-takers to rapidly match individuals of different demographics to words and pictures of varying pleasantness. The quick associations that the test-takers must make reveal their unconscious biases. Additionally, the test-takers often complete a measure of conscious bias. Correlations between scores on the pairing task and the conscious bias scale are generally weak, suggesting that participants have unconscious biases.⁶

Career-related Unconscious Bias

Experimental studies of career-related unconscious bias generally involve participants' rating the skills of equally experienced, fictitious job applicants who vary only by gender or race. As evidenced by the articles described in Table 1, studies of this type support the notion that, given equal weighting on all other variables, people prefer males over females, and white and Asian-American individuals over African-American and Mexican-American individuals in job positions. In contrast to these experi-

mental studies, research in real-life settings highlights the practical ramifications of career-related unconscious bias. Real-life examinations provide evidence that unconscious bias affects decisions in actual hiring and evaluation processes. The articles displayed in Table 2 suggest that unconscious bias negatively affects female and black individuals in search and selection processes. Therefore, employers may be more likely to hire white males over individuals of other demographics for job openings. In contrast to studies involving the IAT, the studies in Tables 1 and 2, which do not address the level of consciousness of the participants' biases, provide only indirect evidence of unconscious bias. However, most individuals believe that they hold fewer biases than the average person,⁷ so the findings of bias in these studies very likely demonstrate unconscious bias.

Implications for Faculty Recruitment Processes in Academic Medicine

As evidenced by the preceding examples, unconscious biases may affect evaluations and decisions made in recruitment processes in academic medicine.

However, the literature suggests that those involved in hiring processes can take steps to mitigate some of the effects of unconscious bias in evaluating job candidates. These steps may include:

- Ask individuals involved in hiring processes to reflect on unconscious biases by taking the online version of the IAT (<https://implicit.harvard.edu/>).
- Create a more objective, structured interview process by (a) setting criteria or using objective measures to assess the skills needed for effective job performance; (b) administering training to all involved in the search and interview process on how to conduct structured interviews; and (c) using performance, satisfaction, and turnover rates of new hires to measure the effectiveness of the interview process.⁸
- Be aware that recommenders of applicants may hold unconscious biases, and therefore may present skewed representations of applicants in their letters of recommendation.⁹
- Consider that cultural differences can affect first impressions of candidates. For instance, the standard interview in the United States uses the criteria of self-confidence, goal orientation, enthusiasm, and leadership, even though these qualities may not be apparent in people of more reserved cultures.¹⁰
- Reserve ample time for the interviews and evaluations of candidates, as sex bias emerges more when evaluators are under time pressure.¹¹

Author:

April Corrice, Research Analyst,
Organizational and Management Studies,
Office of the Executive Vice President,
AAMC, acorrice@aamc.org, 202-741-0747.

Association of American Medical Colleges

2450 N Street, N.W.
Washington, D.C. 20037-1127
analysis@aamc.org
www.aamc.org/data/aib