

Gender Bias and Peer Review: *Annals* Seeks Greater Diversity

Michael Callaham, MD*

*Corresponding Author. E-mail: Michael.Callaham@ucsf.edu.



0196-0644/\$-see front matter

Copyright © 2019 by the American College of Emergency Physicians.

<https://doi.org/10.1016/j.annemergmed.2019.10.018>

SEE RELATED ARTICLE, P. 731.

[Ann Emerg Med. 2019;74:736-739.]

For decades, there has been a growing recognition throughout our society, particularly in science and technology academic fields, that the distribution of gender, cultural, and other forms of diversity in our ranks does not reflect the full population from which experts might be developed. This inequity deprives us of the complete spectrum of talent potentially available. Although we would like this editorial to address all forms of diversity, we will focus on gender because it is the only factor for which we have enough substantive data to draw conclusions. We hope increased awareness, concern, and remediation about one form of inequity will spread to others.

Much research during many years has demonstrated a gender imbalance in science. These studies found women underrepresented in almost all scientific disciplines, including the entire field of medicine, as well as our own specialty of emergency medicine.¹ Academic accomplishments reflecting this bias have included awards, grant funding and other research support, authorship of scientific articles (including the important first and last author positions), presentation of research at meetings, and the roles of editors and peer reviewers who validate and disseminate research, to name a few.² Older studies were recently confirmed by a very large and comprehensive study of 9,000 editors from 142 different journals in multiple disciplines, with 43,000 reviewers of submissions from 126,000 authors.² These disparities have been documented in most countries, and in the most elite research and educational institutions.³ Results in individual scientific disciplines show a great deal of variance. Improvement is slowly (but not steadily) occurring, yet at present rates of change, gender parity would be reached only after 7 years for authors, 14 years for peer reviewers, and 22 years for decision editors.²

The previous research has limitations, of course. Almost all of the studies were retrospective reviews of association

(not causation), and very few have studied physicians (compared with research scientists in “hard” disciplines). Most of these articles did not address any specific remedies. But they certainly are convincing that there is a deep, widespread, and persistent problem that has resulted in women’s achieving fewer academic goals and metrics than men do.

Where does this inequity start? Is it all explained by implicit bias, which is common as early as childhood? Recent studies continue to support the conclusion that women’s actual performance in science is comparable to or higher than men’s in many variables, including high scores on entrance examinations, opportunities to interview for (and be offered) high-quality jobs, salary, successful grant awards, and academic promotions. Yet women’s differing realities about work-life balance and parental roles often drive different priorities (frequently only partially voluntary) in career and professional activities. As just one example, 40% of female graduate students report that because of career pressure, they had fewer children than they wanted; the comparable figure for men is half that. In recent years, the disadvantage of women’s disproportionate childbearing and parenting roles has also been more deeply studied, and has been argued to be as significant as or more significant than other obstacles.⁴ As a leading investigator states, “Academe demands that women having children make their greatest intellectual contributions [eg, achieving tenure]...contemporaneously with their greatest physical and emotional achievements, a feat not expected of men.”⁴ Another large and high-quality recent study, although only descriptive, has cast further light on homophily (the universal human tendency to prefer those similar to ourselves).² This study was able to more precisely analyze the influence of implicit bias in specific scenarios; for example, the fact that although homophily is observed in women (toward other women) as well as men, it is far more common and stronger in male than female journal editors.

In any event, this problem is complex, chronic, and important, and we want to be ahead of it. Our recent

experience at *Annals* is summarized in the detailed report by Kaji et al,⁵ which was discussed at the editorial board's annual retreat in February 2019. Like many other organizations (inside and outside science), we have made progress in some areas, but have much left to do. Table 1 compares our proportion of women in various roles with those in other organizations and environments. Positive aspects are that the proportion of women in various roles at *Annals* has increased, that our resident fellow program has been very successful, and that women are moving into more senior leadership positions. The bad news is that this is not happening fast enough.

Our first formal actions in improving diversity started in 1998 with the creation of our resident fellow position, which is awarded once a year to an applicant currently enrolled in an emergency medicine residency in the United States.⁶ The selected applicant is mentored for a year by our most experienced editors through various peer-review and editorial tasks. The successful applicants (selected by the deputy editors and editor in chief, all men) were 51% women, whereas our editorial board for many years averaged 16% women. This program was not in fact aimed at correcting gender imbalance (a topic seldom discussed in academia or journals at the time) but at recruiting and mentoring motivated young academic physicians of diverse backgrounds with a special interest in scientific publication

Table 1. Comparison of proportion of women in various emergency medicine activities (various dates but the most recent results).

Task	% Women
Resident fellows	58
All medical students	45
Emergency medicine residents	38
<i>Annals</i> core reviewers	31
<i>Annals</i> first authors	31
PLoS journals research authors	30*
<i>Annals</i> original research authors	28
ACEP committee members	28
ACEP council members	27
Regular ACEP members	26
<i>Annals</i> editorial board	24
Invited editorials, authors	21
<i>Annals</i> deputy editors	20 [†]
Editorial board's top 5 emergency medicine journals	17
ACEP senior leadership	12.5

PLoS, *Public Library of Science*; ACEP, American College of Emergency Physicians. ACEP data are from 2016; *Annals* data are from 2018.

*2008 To 2013 PLoS; all data are from *Annals* unless noted otherwise.

[†]Twenty-five percent, including editor in chief.

to benefit and support the future of our journal and our specialty. This effort was very successful and the quality of the selected fellows was very high. About half its participants became permanent reviewers, authors, or both for us. Five of these individuals (3 women) serve on the masthead as editors to this day. Some very talented female fellows went on to even greater leadership beyond our journal: one is now director of the National Center for Injury Prevention and Control at the Centers for Disease Control and Prevention, and another leads the development of the emergency and trauma care program at the World Health Organization in Geneva.

As Braddock⁷ emphasized, concepts for encouraging change and improvement in any activity are not worth much if they are not actionable. Factors that encourage persons to address diversity issues are probably individual, variable, and often difficult to implement. Our optimal goal for *Annals* is to be able to recommend to all our reviewers and editors methods and resources they could use to help recognize and diminish their own (and others') intrinsic bias. The Harvard Implicit Association Test is the best-known tool to claim such an ability. Unfortunately, a recent very comprehensive meta-analysis of 492 studies examining the effectiveness of procedures in changing implicit measures found profound methodological limitations to almost all of them.⁸ Few studies were controlled, the reported effect sizes were very small, and the study populations were overwhelmingly young white women enrolled in college psychology courses. Most studies were also short, usually a few days or even hours, and those that reported any change in bias score generally reverted to the preintervention state in a short period.⁸ Virtually none of these studies would pass even the most generous methodology/statistics review at *Annals*. Therefore, the Implicit Association Test is not accurate enough to guide us to effective actions, a conclusion also reached by its creators.

Two studies deserve special comment because they measured important outcomes and are rare exceptions to the weak methodology of the majority. One examined the actions of real researchers in grant review committees who were making real decisions about real applicants.⁹ Also uniquely, it had a control group. Limitations included that subjects were researchers from only one country (France) and were in the basic sciences, not medicine. This study examined the utility of (mostly educational) measures used to reduce existing bias, using sound but complicated methodology that could produce credible evidence for its recommendations. In this study, the measured implicit bias of grant reviewers against women (as manifested by lower rating scores on their grants compared with those of men)

was significantly decreased when reviewers, before assigning scores, received explicit information and discussion of that bias and its effect in the real world.⁹ A second real-world study with a more rigorous study design studied 23,900 Canadian health grant applications reviewed between 2011 and 2016.¹⁰ Ratings were compared when based only on an evaluation of the proposal itself versus evaluation of the investigator submitting the proposal. The former demonstrated no gender bias, whereas when the investigator was the subject of review, men were 1.4 times more likely to receive funding.¹¹ Men may have small systemic advantages (different levels of research support, clinical service load, lesser parental responsibilities, and lack of sexual harassment) that over time would steadily compound into a large systemic advantage. The authors suggested limiting the investigator rating to a binary (capable versus not). Their findings illustrate that more rigorous and accurate assessment tools are strongly needed.

We still do not have the proven detection and remedial tools we would need to sufficiently enhance the success of our efforts. We must forge on without them. With what we know from the research literature and our own experience, we can speculate that some of the following might be actionable. We will explore possible ways to encourage and support women as first or last authors, as well as ask more women to write editorials or other invited publications. We should make sure high-performing reviewers (especially women) are aware of the fact that the pathway to editorship begins with a strong profile as a peer reviewer. We can encourage even more engagement of editors with reviewers (especially new ones) by assigning more reviewers per article, particularly to underrepresented groups, and

making this a priority for our editors. Wherever possible, we should encourage more interaction and mentoring between reviewers and decision editors, and more frequent reminders of our equity goals.

For decades we have closely assessed and curated our reviewer pool every year, ranking them for availability, volume, promptness, and quality of reviews. Reviewers receive continuing medical education credit for their work and summary letters of their reviewer accomplishments in support of academic promotions. Good reviewers are crucial to our success, and we try to thank them appropriately by identifying them publicly. An impressive number of them qualify for our senior reviewer status (top 50 highest performers at least twice in 4 years) and distinguished senior reviewers (qualified as senior reviewers at least 6 of the last 10 years). This not only honors their contributions to the journal but also provides us a pool of experts who have already proven their high competency in assessing and improving science. Thus, when the occasional editor position opens up, these lists are a natural source of candidates for the position. Unfortunately, the proportion of women in these groups is profoundly low, and we are still trying to determine the cause. We have also placed a special priority on finding strong female candidates for more senior editor positions because that is where the deficit is greatest; recently, we have filled a deputy editor position with a very experienced female editor. We will expand this process whenever possible. We will also try to fill the pipelines with strong female talent, knowing that although they may not have many years of publishing experience now, they are the future editors and editorial leadership of this journal. Additionally, knowing that

Table 2. Educational and reference resources in peer review.

Annals Journal Club	https://www.annemergmed.com/journalclub	Examples of critical analysis
COPE (excellent algorithms)	https://publicationethics.org	Sets best practices, also problematic cases with expert commentary
Scholarly Kitchen	https://scholarlykitchen.sspnet.org	Blog on scientific publication
Original research on scientific publication	Perform and publish research relevant to peer review yourself	
WAME	Discussions, feedback from other editors, develops education for editors (WAME.org)	Sets best practices
Peer Review Congress (<i>JAMA</i> , <i>BMJ</i>) Researchers meet and discuss presentations	https://peerreviewcongress.org/program-information (original research on scientific publishing is presented)	Every 5 years, new research and editorial panels and discussions
Peer-review training module, software	https://els-jbs-prod-cdn.literatumonline.com/pb/assets/raw/Health%20Advance/journals/ymem/index-1539617262570.html	Educational software examining the peer-review process
Implicit bias test (unconscious bias)	https://implicit.harvard.edu/implicit/	The actual test; authors warn that it has not been validated.

COPE, Council on Publication Ethics; WAME, World Association of Medical Editors.

homophily is universal in human behavior, we will try to develop mechanisms to help editors integrate thinking about gender bias into every relevant decision. Finally, to keep ourselves honest, we will assess and report our progress on diversity annually.

What should a physician interested in participating more in the world of peer review publishing do at *Annals* to achieve that goal? The first step is become a peer reviewer; a training module and other resources are listed in Table 2. Developing a reputation for expertise in an academic field of research is very helpful. While writing your reviews, do not hesitate to ask your decision editor or experienced colleagues for discussion of difficult issues and ways in which your reviews could be improved. Compare your finished review with others on the same article to see what strengths and weaknesses they addressed compared with yours. *Annals* Journal Clubs allow you to compare your review of a real article with key issues mentioned by experienced discussants. The Council on Publication Ethics and World Association of Medical Editors are organizations that set standards in scientific publication (see URLs in Table 2). When you accumulate substantial reviewer and researcher experience, share your achievements with one of the deputy editors if you are seeking a future editorial position.

Author affiliations: From the Department of Emergency Medicine, University of California, San Francisco, San Francisco, CA.

Authorship: All authors attest to meeting the four ICMJE.org authorship criteria: (1) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND (2) Drafting the work or revising it critically for important intellectual content; AND (3) Final approval of the version to be published; AND (4) Agreement to be accountable for all aspects of the work in ensuring that questions

related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding and support: By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see www.icmje.org). The author reports being the editor of *Annals of Emergency Medicine* and being paid a stipend for that purpose by the American College of Emergency Physicians.

REFERENCES

1. Parker RB, Stack SJ, Schneider SM, et al. Why diversity and inclusion are critical to the American College of Emergency Physicians' future success. *Ann Emerg Med*. 2017;69:714-717.
2. Helmer M, Schottdorf M, Neef A, et al. Gender bias in scholarly peer review. *Elife*. 2017;6:1-18.
3. Larivière V, Ni C, Gingras Y, et al. Global gender disparities in science. *Nat News*. 2013;504:211-213.
4. Ceci SJ, Williams WM. Understanding current causes of women's underrepresentation in science. *Proc Natl Acad Sci U S A*. 2011;108:3157-3162.
5. Kaji AH, Meurer WJ, Napper T, et al. State of the journal: women first authors, peer reviewers, and editorial board members at *Annals of Emergency Medicine*. *Ann Emerg Med*. 2019;74:731-735.
6. Reynolds TATA, Callaham M. A survey of past participants in the *Annals of Emergency Medicine* editorial board fellowship program. *Ann Emerg Med*. 2011;57:161-164.
7. Braddock CH III. Strategies to enhance diversity in emergency medicine: finding actionable solutions. *Ann Emerg Med*. 2019;74:751-752.
8. Forscher PS, Lai CK, Axt JR, et al. A meta-analysis of procedures to change implicit measures. *J Pers Soc Psychol*. 2019;117:522-559.
9. Régner I, Thinus-Blanc C, Netter A, et al. Committees with implicit biases promote fewer women when they do not believe gender bias exists. *Nat Hum Behav*. 2019; <https://doi.org/10.1038/s41562-019-0686-3>.
10. Witteman C, Harries C, Bekker H, et al. Evaluating psychodiagnostic decisions. *J Eval Clin Pract*. 2007;13:10-15.
11. Raymond JL, Goodman MB. Funders should evaluate projects, not people. *Lancet*. 2019;393:494-495.

Did you know?

You can track the impact of your article with citation alerts that let you know when your article (or any article you'd like to track) has been cited by another Elsevier-published journal.

Visit www.annemergmed.com today to see what else is new online!