Supporting Information

Ceci and Williams 10.1073/pnas.1014871108

SI Text

S1. Consider the percentages of full professors: computer science, 10.3%; chemistry, 9.7%; economics, 8.7%; mathematics, 7.1%; physics, 6.1%; and mechanical engineering, 4.4% (1). Generally, doctorate-granting institutions employ a lower percentage of women on tenure track than do liberal arts colleges. For example, in economics, the percentages, collapsed across ranks, are 16.7% and 27.6%, respectively (2). Percentages of women in feeder pipelines in these math-intensive fields are somewhat higher than in these combined ranks, but generally track the assistant professor percentages. In 2005, the percentages of PhDs awarded to women in highly quantitative fields were as follows: chemistry, 32.4%; mathematics, 30%; computer science, 21% (3); physics, 14.3%; chemical engineering, 23.7%; civil engineering, 22%; electrical engineering, 12.3%; and mechanical engineering, 8.4% (1). Percentages of women hired on tenure track were as follows: chemistry, 21.2%; mathematics, 26.8%; computer science, 20.0%; physics, 16.8%; chemical engineering, 24.2%; civil engineering, 24.7%; electrical engineering, 15.5%; and mechanical engineering, 18.0%. As can be seen, in some fields the percentage of women with a Ph.D. is larger than the percentage of women hired on tenure-track (e.g., chemistry), and in other fields, the reverse is true (e.g., mechanical engineering), but there is no systematic over- or underrepresentation.

S2. In the debate about the causes of sex inequality in math-intensive fields of science, some argue that whenever a given hiring outcome falls outside a preordained threshold, this is evidence of discrimination. For example, in 2006, of all PhDs awarded in mathematics, 29.6% went to women (3). However, women comprised a significantly smaller percentage of "tenure eligible" assistant professors hired that year, and they have comprised a smaller percentage of all tenured and tenure-eligible posts in mathematics departments in the past half century than is reflected by their pipeline numbers. Below we provide the evidence for these statements.

Women's entry into tenure-track mathematics departments has been growing in recent years but still falls short of their representation in the PhD pipeline. For example, table F1 of the Conference Board of the Mathematical Sciences Survey (4, p. 103) lists women faculty as comprising 218 tenure-eligible faculty (out of a total of 930) at PhD-granting departments, totaling 23.5% of all "tenure-eligible" faculty. The same table lists 420 tenured female faculty of a total of 4,699 tenured faculty. Summing these two figures results in 5,629 total tenured and tenure-eligible faculty, of which only 638 (11%) are women. [This is similar to table 11 of Nelson and Brammer (1) at 12.9% women, and given that their data were two years newer than the CBMS survey (4) and were based on the top 100 departments of mathematics as opposed to all PhD-granting departments, these sets of figures concord.] Between 1958 and 2006, women received 17.3% of all PhDs awarded in mathematics (3, table 1). So, we can assume that women in tenured and tenure-eligible posts are represented significantly below their recent representation in PhD programs (29.6% in 2006) as well as below their historic representation (17.3%). What we do not know is why.

Our article suggests that reasons for women's underrepresentation may have changed. Although unfair biases based on invalid stereotypes, and gender bias in hiring and remuneration may have played a significant role in the dearth of women scientists in the past, other factors are required to explain this phenomenon today. Referring to current factors as "discrimination" is not only inaccurate, it also leads to interventions that are unlikely to remedy the underrepresentation (e.g., requiring search committee members and grant panel reviewers to be certified in "gender sensitivity training"). According to some, the gap between the potential pool of female applicants in PhD programs and the number of women actually hired on tenure track is prima facie evidence of discrimination on the basis of sex. However, this definition of discrimination presupposes many factors that have not been demonstrated to be true and, indeed, may not be true: Do women apply for the same jobs over as wide a geographical area as men? Are women as motivated to attain and accept the jobs (if offered) as men? Are women's publications and records equivalent to men's? Are the women as likely as the men to regard tenure track positions at research-intensive universities as family-friendly and compatible with their plans?

Again, for discrimination to be shown to exist, the women and men must possess similar credentials and must apply equally vigorously and be equally willing to relocate. If such aspects of the situation are true and can be verified, and if women are passed over for jobs relative to men, then discrimination is taking place. However, if women choose not to apply, won't relocate, and/or lack the scholarly records possessed by their male counterparts, we cannot infer discrimination, because the characteristics of the applicants and their situations are not equivalent.

In sum, simple outcome differences between two groups do not prove that members of the underrepresented group are victims of discrimination. We know from the 2009 NRC analysis that women in math-intensive fields who apply for tenure-track positions are invited to interview and are hired at slightly higher rates than their male colleagues (6), thus there is no evidence that search committees pass over female candidates. The problem is that many more women than men opt out of applying for these positions in the first place. When women PhD recipients choose not to apply for tenure-track posts, their refusal represents a choice, one that most of their male and many of their female colleagues do not make.

From GAO exit interviews (7), it seems that at least some of women's choice not to pursue tenure-track positions results from deciding that a career in academia is incompatible with family formation and work-life balance. A great deal of survey evidence amply makes the case that women on tenure track are faced with difficult challenges—ones not faced by men—and this is a top reason given by female PhD candidates in science for switching out of science. As we noted in footnote 3 of our article, analyses of >8,000 University of California graduate students' responses by Mason and Goulden (8) document the role played by fertility decisions in female graduate students' switching out of tenure-track career aspirations in science. Married women doctoral students with children are 35% less likely to enter a tenure-track position after receiving a Ph.D. than are married men with children, and women are 27% less likely than men to achieve tenure, whereas single females without children achieve comparably to their male counterparts (5, 8, 9). Puuska showed that it is the presence of young children (as opposed to older children) that is associated with lower productivity of female faculty (32).

We wish to emphasize the substantial challenges faced uniquely by women, resulting from the differing biological realities of the sexes and the lifecourse needs and decisions eventuating from these differing biological realities. However, to call differential outcomes resulting from women's choices "discrimination" is misleading, because doing so conflates consequences of choices, both freely made and constrained (e.g., to have children, to follow a partner, to care for elderly parents, to assume a disproportionate role in childcare), with the practice of making invalid distinctions between people who are otherwise similar in all aspects relevant to some criterion, based on their group membership (e.g., race, sex, social class, sexual orientation). We reserve use of the label of sex discrimination for cases in which equally qualified female candidates are not hired because of their sex, a practice which the data show no longer takes place in the academy.

Universities can and should do more to ease female faculty members' burdens, through the types of programs that sex equity committees, governmental agencies, and others have proposed (see main article for examples; see also refs. 8 and 9 for California's Family Friendly recommendations). However, little is to be gained by labeling as "discrimination" outcomes of decisions some women make that other women do not. A more accurate description would be that women, in part because of biological realities surrounding pregnancy, birth, and motherhood and, in part, because of gendered expectations resulting in women assuming far more family responsibilities than men do, are confronted with often-overwhelming demands not faced by men. Misuse of the term discrimination obfuscates the real issues women face and delays or prevents progress toward resolving them, because the remedies for hiring discrimination are very different from the remedies for pipeline issues that impede women's transition from graduate school to the academy and from junior to senior ranks within the academy.

S3. One reviewer noted the "chicken and egg" aspect of this point: "the notion that differences in acceptance rates disappear once support structures covarying with institutional quality are taken into account is important. However, is this due to the support structure or to the fact that brighter people end up in better places? One might argue (and I certainly would) that due to pronounced preferences for females in academic hiring (10), junior men in high quality institutions have to be slightly more gifted than their female counterparts. Consequently, the fact that there is no difference in acceptance rates probably indicates a slight bias in favor of female submissions." In contrast to this data-based view is the assertion that parity in grant, hiring, and manuscript decisions is misleading, because women are superior to men, and therefore ought not to be at parity with, but rather, above men. On this view, the evidence that women publish as first authors equivalently to men (e.g., Cortex), receive grants at comparable rates, or get hired at comparable or even slightly higher rates than men is insufficient to prove sex bias has ceased. Such assertions regarding superior quality of women are unfalsifiable with available data. The most relevant, albeit imperfect, data suggest male and female scientists' papers are of comparable quality; e.g., researchers found no differences in citation rates of male and female papers (11). This lack of difference suggests the reviewing process is unbiased in relation to sex; otherwise, if the smaller number of women were of higher quality, their citations should have been more numerous than men's.

S4. With such small samples, slight adjustments can matter. For instance, what if the review panels were disproportionately composed of biomedical (not nursing or basic science) professionals and more male applicants came from medical or biomedical backgrounds? Only 27% of female applicants came from medical backgrounds vs. 60% of males. One need not posit sex discrimination if reviewers preferred medical journals to basic science or nursing journals, even if the latter were more highly cited. Reviewers were not given impact ratings of the journals or other citation information and may have relied on their judgments about which fields are stronger, even if these fields are less populated than weaker fields and, consequently, engender fewer citations. If such field-based preferences existed, the greater proportion of female applicants from nursing (12% female applicants vs. 3% male applicants) could have tilted the odds against them, even if nursing journals were cited more frequently. In other words, quantifying applicants' quality by adjusting applicants' publications by citation factors may have resulted in an objective measure that differed from reviewers' judgments of quality of these same publications.

Furthermore, some have criticized the regression models described in Wennerås and Wold's article, because the authors entered each productivity variable alone into their equations rather than allowing for multiple variables to enter, as might be expected in the real world as one commentator noted: "I sent the Swedish study to two research psychologists, Jerre Levy (professor emerita, University of Chicago) and James Steiger (professor and director, Quantitative Methods and Evaluation, Department of Psychology and Human Development, Vanderbilt) for their review. They both immediately zeroed in on a troubling methodological anomaly: Wennerås and Wold had run separate regressions for only one productivity variable at a time. Since it is unlikely that any single variable adequately characterizes academic productivity, the obvious approach would have been to enter several of the productivity variables into a single regression equation… Steiger wrote to Wennerås and Wold requesting copies of the data so he could review them himself. Wold wrote back that she would gladly send the data, except that they had gone missing" (p. 91, 12). Others have questioned the authors' assumption of linearity, e.g., the difference between 1 and 2 publications may not be tantamount to the difference between 14 and 15. Also, it is possible to reverse the causality of Wennerås and Wold's model and argue that men had to be 2.6 times as competent to be published in journals as women and 10 times as competent to get a fellowship (13). Although this hypothetical account is unlikely to be true, this possibility cannot be ruled out. The causal logic between the variables used by Wennerås and Wold can operate in either direction, and possibilities such as sheer number of journal publications being as biased against males as grant reviewers' ratings of scientific competence are biased against females cannot be refuted.

Finally, often when male and female scientists' productivity is ranked from highest to lowest, males are overrepresented at the high end, usually around 2-to-1. For example, in a national study of young Croatian scientists, it was reported that the proportion of highly productive young male scientists is greater than that of females: "24.3% men and 12.2% women published $\geq$10 articles, while 10.5% men and 4.1% women published $\geq$15 articles in a five-year period" (14, p. 40). Similarly, in a population study of the top 10% of Italian scientific productivity, 8% of males had an average publication rate of $\geq$10 articles per year, vs. only 2% of females (16). If the Swedish reviewers were faced with the reality that they could only fund $\approx$18% of applicants (20 of 114), and if the most productive group contained a 2:1 ratio of males to females, then this would have resulted in roughly 14 males and 7 females being funded, which is not reliably different statistically from what they did fund (16 vs. 4, $P > 20$). Such an outcome remains possible even if the total impact score was most predictive of the reviewers' ratings of scientific competence, $r^2 = 0.47$, because of this presumed sex asymmetry at the high end.

Efforts to obtain the authors' data for reanalysis of this possibility and to rerun the regression models have been unsuccessful, although one Swedish researcher has reportedly replicated their findings by using the same 1994 data they used. Ultimately, we must await the public disclosure of the original Wennerås and Wold data to test these possibilities, but even if their finding stands, it is clear that the vast majority of research, including larger, more sophisticated studies, do not support their conclusion of sex discrimination in grant funding, including more recent analyses of the Swedish MRC funding data which shows, as noted in our article, a bias in favor of female applicants.

S5. However, RAND did find a sex gap in reapplication rates, with males more likely to apply for follow-up grants (15). Why men were more likely to reapply for subsequent funding or why their grants tended to be funded at higher dollar amounts at NIH (but not at NSF or USDA) is unknown. Because the NIH data did not include the amount of funding requested, it is unclear whether the sex differences reflect differences in the amount males and females request or NIH decisions about how much to award, or both, leading the authors to conclude "If these covariates affect the funding of NIH awards as they do at NSF, the gender gap could be smaller" (p. 3).

S6. In a simulation based on a "typical" hypothetical pre-PhD applicant, averaged across sex, institution, and discipline, Bornmann and Daniel (33) found that the probability of being approved decreased from a male at 50% to a female at 33%. Thus, this study provides some support for approval biases due to sex existing between 25 and 20 y ago, although the authors point out that group differences may be invalid, may be based on invalid characteristics, or may be valid—a distinction requiring an experiment. It is worth noting that this analysis did not find sex differences for the group studied by Wennerås and Wold, postdoc fellowships.

S7. They also found evidence for a sex effect favoring males and effects of major field of study and institutional bias. However, they found these effects only for approval rates of applications for doctoral, but not postdoctoral, fellowships. Thus, they diverge from the findings of Wennerås and Wold of bias at the postdoctoral level.

S8. Using a nationally representative sample of PhD recipients between 1975 and 2000, Ginther found that not all sex differences in salary and promotion disappeared after controlling for citations and publications, particularly among full professors (17, 18). Her coefficients reveal that a productivity increase bought a female full professor in this cohort less salary than a comparable male. However, her later analyses with Kahn (19) concluded that when the appropriate statistical controls are exerted, such as controls for having young children, the evidence of past discrimination disappears, with the sex gap entirely explained by fertility decisions. "We find that in science overall, there is no gender difference in promotion to tenure or to full professor after controlling for demographic, family, employer and productivity covariates and that in many cases, there is no gender difference. . .even without controlling for these covariates" (19, p. 1). Although Ginther and Kahn found historical discriminatory remuneration among full professors, this has not been uniformly found. Levin and Stephan (20) found pay increases in biochemistry, physics, earth science, and physiology in the 1973–1979 period were sex-neutral: "There is little evidence to suggest that the reward process during this period was systematically based on gender. The reward for another publication does not appear to be less for women than for men . . .. (and) in physics women's research productivity was more handsomely rewarded than men's during this period" (p. 1061). This difference of opinion appears to have been resolved during the past decade, with Ginther also finding that a sex difference in pay for comparable jobs in the academy is not a plausible cause of women's lower numbers in math-intensive fields.

S9. In adolescent surveys, it is less common for girls to name math-intensive career aspirations. In one recent poll of 8- to 17-y-olds, 24% of boys expressed interest in engineering vs. only 5% of girls; in a survey of 13- to 17-y-olds, 74% of boys expressed interest in computer science vs. only 32% of girls. Large sex differences are evident in the people-versus-things dimension of vocational interests, according to a recent metaanalysis (see ref. 21, p. 4).

S10. Although there are no systematic sex differences at the midpoint of the mathematics achievement distribution, there are fairly large differences at the extreme tails, both right and left. "Among the top 1% of students on standardized mathematics tests, there are approximately 2 males for every female, and this ratio has been found across a wide variety of nationally representative samples (e.g., ref. 22, n = 7 million US students in various grades; ref. 23, n = 318,599 9- to 11-year-olds; ref. 24, n = 0.5 million; ref. 25, n = 1.6 million 7th graders). Because overall male score variability is roughly 0.15 SD greater than females' on these mathematics tests, the farther out on either tail one looks, the higher is the ratio of males to females. During the 2006–2010 period, in a nonrandom sample of 7th grade perfect-800 scorers on the SAT-Mathematics, there were 6.58 males for every female (26). In the past 20 years, there have been 37 perfect scorers among 7th graders on the ACT-Science test, 36 of whom were male. Lest these data be assumed to reflect innate, biological superiority of males, it deserves noting that there is large variance in cross-cultural analyses, with females outperforming males at the right tail in some countries (27, 28), and the best predictor of international sex differences is the degree to which its citizens exhibit implicit gender-science stereotypes (29) (see ref. 30 for references). Finally, somewhat offsetting the male overrepresentation at the right tail on math gateway tests is the fact that females earn better grades in mathematics, including in college courses, thus reducing any male advantage on graduate school admissions based solely on GRE-Q scores.

S11. The GAO report (7) noted that "Women PhD students we interviewed revealed that very few would seek tenure track positions at research institutions. Most said that they would rather become faculty at small colleges or scientists at a laboratory where they thought work pressures would be less intense and they could maintain a more healthy balance between work and family life" (p. 20). In comments posted online in response to a column in the *New York Times*' "Science Times" (31), poster 20 commented: "I'm currently a Ph.D. student in neuroscience. If I'm very lucky, I'll have tenure by the time I'm 35. I don't feel it's realistic to wait to have children until I'm that old. But what am I supposed to do? Try to raise a child on the pathetic salary of a post-doc while working 60 (or more) hours a week? That is the decision many, many women in science now face. And it's not like you can take a break and come back to science. Once you're out, it's next to impossible to get back in. I can't believe someone would claim that being an academic scientist isn't a huge barrier to having a family. Almost all of my female role models (of which there are relatively few) do not have kids. One very well known female scientist even once told a group of us that her biggest regret is that she didn't freeze some of her eggs so that she could have kids later in life (by the time she felt she was in a position to try career-wise, she was infertile). True, men confront these problems as well. But the reality is, men stay fertile much later in life, and they often have wives who do most or all of the child-care for them. Men also still tend to earn more money than women, meaning there's more pressure for the woman to be the one to take time off for the kids. If (another poster) is wondering why there are more women teachers, may I suggest it has something to do with the fact that you are only at work while your kids are at school? And that you can enter the field later in life, after being pushed out of another career due to the fact you decided to take time off while your kids were pre-K?"

The number of young women who want to pursue careers in academic research declines by 30% over the course of their doctoral study, and the number of men by 20% (8, 9). Thus, 50% more women than men opt out at this stage. In explaining their decision, men are more likely to report they do not like the unrelenting work hours. One male student in the survey complained that he was "fed up with the narrow-mindedness of supposedly intelligent people who are largely workaholic and expect others to be so as well." However, more women give up on academic-research careers for family concerns. As one woman in the survey said, "I could not have come to graduate school more motivated to be a research-oriented professor. Now I feel that can only be a career possibility if I am willing to sacrifice having children."

1. Nelson D, Brammer C (2010) A National Analysis of Minorities in Science and Engineering Faculties at Research Universities. Available at http://chem.ou.edu/~djn/diversity/faculty_Tables_FY07/FinalReport07.html. Accessed June 17, 2010.
2. Fraumeni BM (2008) Report of the Committee on the Status of Women in the Economics Profession 2008. Available at http://www.aeaweb.org/committees/cswep/annual_reports/2008_CSWEP_Annual_Report.pdf. Accessed January 12, 2011.
3. Burrelli J (2008) *Thirty-three years of women in S&E faculty positions*. InfoBrief NSF 08-308. (National Science Foundation, Division of Science Resources Statistics, Arlington, VA).
4. Lutzer DJ, Rodi SB, Kirkman EE, Maxwell JW (2007) Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2005 survey. *Amer Math Soc*, http://www.ams.org/profession/data/cbms-survey/cbms2005. Accessed January 12, 2011.
5. Mason MA, Goulden M (2004) Marriage and baby blues: Redefining gender equity and the academy. *Ann Amer Polit and Soc Sci* 596:86–103.
6. Faculty Committee on Women in Science, Engineering, and Medicine (2010) *Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty.* (Natl Acad Press, Washington DC).
7. General Accounting Office (July 2004) *Gender issues: Women's participation in the sciences has increased, but agencies need to do more to ensure compliance with title IX. GAO 04-639*, Available at www.gao.gov/cgi-bin/getrpt?GAO-04-639.
8. Mason MA, Goulden M (2009) UC Doctoral Student Career and Life Survey. (Univ of Calif, Berkeley). Available at http://ucfamilyedge.berkeley.edu/grad%20life%20survey.html. Accessed January 5, 2011.
9. Mason MA, Goulden M, Frasch K (2009) Why graduate students reject the fast track. *Academe* 95:11–16.
10. Irvine AD (1996) Jack and Jill and employment equity. *Dialogue* 35:255–291.
11. Tregenza T (2002) Gender bias in the refereeing process? *Trends Ecol Evol* 17: 349–350.
12. Sommers CH (2008) Why can't a woman be more like a man? The American. Available at http://www.american.com/archive/2008/march-april-magazine-contents/why-can2019t-a-woman-be-more-like-a-man/?searchterm=christina%20sommers. Accessed January 13, 2011.
13. Hansson O (2009, March) Wold and Wenneras solve 2 unknown variables with 1 equation and receive honorary degree from Chalmers. Available at http://www.athleticdesign.se/otherstuff/sexism_in_peer_review_english.html. Accessed June 28, 2010.
14. Prpic K (2002) Gender and productivity differentials in science. *Scientometr* 55:27–58.
15. RAND (2005) Is there gender bias in federal grant programs? (RAND Infrastructure, Safety, and Environment Brief RB-9147-NSF). Available at http://rand.org/pubs/research_briefs/RB9147/RAND_RB9147.pdf. Accessed March 1, 2008.
16. Abramo G, D'Angelo C, Capraesecca A (2009) The contribution of star scientists to overall sex differences in research productivity. *Scientometr* 81:137–156.
17. Wennerås C, Wold A (1997) Nepotism and sexism in peer-review. *Nature* 387:341–343.
18. Ginther DK (2001, February) Does Science Discriminate against Women? (Fed Res Bank of Atlanta, Atlanta). Working Paper 2001-02.
19. Ginther DK, Kahn S (2006) Does Science Promote Women? Science and Engineering Careers in the United States: An Analysis of Markets and Employment, eds Freeman RB, Goroff DL (Natl Bur Econ Res, Chicago) pp 163–194.
20. Levin S, Stephan P (1998) Gender differences in rewards to publishing in academe: Science in the 1970s. *Sex Roles* 38:1049–1064.
21. Su R, Rounds J, Armstrong P (2009) Men and things, women and people: A meta-analysis of sex differences in interests. *Psych Bull* 135:859–884.
22. Hyde JS, Lindberg SM, Linn MC, Ellis AB, Williams CC (2008) Diversity. Gender similarities characterize math performance. *Science* 321:494–495.
23. Lohman DF, Lakin JM (2009) Consistencies in sex differences on the Cognitive Abilities Test across countries, frades, test forms, and cohorts. *Br J Educ Psychol* 79: 389–407.
24. Strand S, Deary IJ, Smith P (2006) Sex differences in cognitive abilities test scores: A UK national picture. *Br J Educ Psychol* 76:463–480.
25. Mullis IVS, Martin MO, Fierros EG, Goldberg AL (2000) *Gender differences in achievement: IEA's Third International Mathematics and Science Study* Boston College International Study Center, Chestnut Hill, MA), timss.bc.edu/timss1995i/TIMSSPDF/T95_GChapter%201.pdf. Accessed May 1, 2008.
26. Wai J, Cacchio M, Putallaz M, Makel M (2010) Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence* 38:412–423.
27. Ceci SJ, Williams WM, Barnett SM (2009) Women's underrepresentation in science: Sociocultural and biological considerations. *Psychol Bull* 135:218–261.
28. Else-Quest NM, Hyde JS, Linn MC (2004) Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol Bull* 136:103–127.
29. Nosek BA, et al. (2008) National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proc Natl Acad Sci USA* 106: 10593–10597.
30. Valla J, Ceci S, et al. Can sex differences in science be tied to the long reach of prenatal hormones? *Perspect Psychol Sci*, in press.
31. Tierney J (June 14, 2010) Legislation won't close gender gap in science. *New York Times*, p D4. Available at http://www.nytimes.com/2010/06/15/science/15tier.html?_r=1&scp=1&sq=stephen%20ceci&st=cse. Accessed January, 13, 2011.
32. Puuska 1 H-M (2010) Effects of scholar's gender and professional position on publishing productivity in different publication types. Analysis of a Finnish university. *Scientometrics* 82:419–437.
33. Bornmann L, Daniel H.-D. (2005) Selection of research fellowship recipients by committee peer review. *Scientometr* 63:297–320.