Science faculty's subtle gender biases favor male students

Corinne A. Moss-Racusin[a,b], John F. Dovidio[b], Victoria L. Brescoll[c], Mark J. Graham[a,d], and Jo Handelsman[a,1]

[a]Department of Molecular, Cellular and Developmental Biology, [b]Department of Psychology, [c]School of Management, and [d]Department of Psychiatry, Yale University, New Haven, CT 06520

Edited* by Shirley Tilghman, Princeton University, Princeton, NJ, and approved August 21, 2012 (received for review July 2, 2012)

Despite efforts to recruit and retain more women, a stark gender disparity persists within academic science. Abundant research has demonstrated gender bias in many demographic groups, but has yet to experimentally investigate whether science faculty exhibit a bias against female students that could contribute to the gender disparity in academic science. In a randomized double-blind study (*n* = 127), science faculty from research-intensive universities rated the application materials of a student—who was randomly assigned either a male or female name—for a laboratory manager position. Faculty participants rated the male applicant as significantly more competent and hireable than the (identical) female applicant. These participants also selected a higher starting salary and offered more career mentoring to the male applicant. The gender of the faculty participants did not affect responses, such that female and male faculty were equally likely to exhibit bias against the female student. Mediation analyses indicated that the female student was less likely to be hired because she was viewed as less competent. We also assessed faculty participants' preexisting subtle bias against women using a standard instrument and found that preexisting subtle bias against women played a moderating role, such that subtle bias against women was associated with less support for the female student, but was unrelated to reactions to the male student. These results suggest that interventions addressing faculty gender bias might advance the goal of increasing the participation of women in science.

diversity | lifestyle choices | science education | science workforce

A 2012 report from the President's Council of Advisors on Science and Technology indicates that training scientists and engineers at current rates will result in a deficit of 1,000,000 workers to meet United States workforce demands over the next decade (1). To help close this formidable gap, the report calls for the increased training and retention of women, who are starkly underrepresented within many fields of science, especially among the professoriate (2–4). Although the proportion of science degrees granted to women has increased (5), there is a persistent disparity between the number of women receiving PhDs and those hired as junior faculty (1–4). This gap suggests that the problem will not resolve itself solely by more generations of women moving through the academic pipeline but that instead, women's advancement within academic science may be actively impeded.

With evidence suggesting that biological sex differences in inherent aptitude for math and science are small or nonexistent (6–8), the efforts of many researchers and academic leaders to identify causes of the science gender disparity have focused instead on the life choices that may compete with women's pursuit of the most demanding positions. Some research suggests that these lifestyle choices (whether free or constrained) likely contribute to the gender imbalance (9–11), but because the majority of these studies are correlational, whether lifestyle factors are solely or primarily responsible remains unclear. Still, some researchers have argued that women's preference for nonscience disciplines and their tendency to take on a disproportionate amount of child- and family-care are the primary causes of the gender disparity in science (9–11), and that it "is not caused by discrimination in these domains" (10). This assertion has received substantial attention and generated significant debate among the scientific community, leading some to conclude that gender discrimination indeed does not exist nor contribute to the gender disparity within academic science (e.g., refs. 12 and 13).

Despite this controversy, experimental research testing for the presence and magnitude of gender discrimination in the biological and physical sciences has yet to be conducted. Although acknowledging that various lifestyle choices likely contribute to the gender imbalance in science (9–11), the present research is unique in investigating whether faculty gender bias exists within academic biological and physical sciences, and whether it might exert an independent effect on the gender disparity as students progress through the pipeline to careers in science. Specifically, the present experiment examined whether, given an equally qualified male and female student, science faculty members would show preferential evaluation and treatment of the male student to work in their laboratory. Although the correlational and related laboratory studies discussed below suggest that such bias is likely (contrary to previous arguments) (9–11), we know of no previous experiments that have tested for faculty bias against female students within academic science.

If faculty express gender biases, we are not suggesting that these biases are intentional or stem from a conscious desire to impede the progress of women in science. Past studies indicate that people's behavior is shaped by implicit or unintended biases, stemming from repeated exposure to pervasive cultural stereotypes (14) that portray women as less competent but simultaneously emphasize their warmth and likeability compared with men (15). Despite significant decreases in overt sexism over the last few decades (particularly among highly educated people) (16), these subtle gender biases are often still held by even the most egalitarian individuals (17), and are exhibited by both men and women (18). Given this body of work, we expected that female faculty would be just as likely as male faculty to express an unintended bias against female undergraduate science students. The fact that these prevalent biases often remain undetected highlights the need for an experimental investigation to determine whether they may be present within academic science and, if so, raise awareness of their potential impact.

Whether these gender biases operate in academic sciences remains an open question. On the one hand, although considerable research demonstrates gender bias in a variety of other domains (19–23), science faculty members may not exhibit this

Author contributions: C.A.M.-R., J.F.D., V.L.B., M.J.G., and J.H. designed research; C.A.M.-R. performed research; C.A.M.-R. analyzed data; and C.A.M.-R., J.F.D., V.L.B., M.J.G., and J.H. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: jo.handelsman@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1211286109/-/DCSupplemental.

bias because they have been rigorously trained to be objective. On the other hand, research demonstrates that people who value their objectivity and fairness are paradoxically particularly likely to fall prey to biases, in part because they are not on guard against subtle bias (24, 25). Thus, by investigating whether science faculty exhibit a bias that could contribute to the gender disparity within the fields of science, technology, engineering, and mathematics (in which objectivity is emphasized), the current study addressed critical theoretical and practical gaps in that it provided an experimental test of faculty discrimination against female students within academic science.

A number of lines of research suggest that such discrimination is likely. Science is robustly male gender-typed (26, 27), resources are inequitably distributed among men and women in many academic science settings (28), some undergraduate women perceive unequal treatment of the genders within science fields (29), and nonexperimental evidence suggests that gender bias is present in other fields (19). Some experimental evidence suggests that even though evaluators report liking women more than men (15), they judge women as less competent than men even when they have identical backgrounds (20). However, these studies used undergraduate students as participants (rather than experienced faculty members), and focused on performance domains outside of academic science, such as completing perceptual tasks (21), writing nonscience articles (22), and being evaluated for a corporate managerial position (23).

Thus, whether aspiring women scientists encounter discrimination from faculty members remains unknown. The formative predoctoral years are a critical window, because students' experiences at this juncture shape both their beliefs about their own abilities and subsequent persistence in science (30, 31). Therefore, we selected this career stage as the focus of the present study because it represents an opportunity to address issues that manifest immediately and also resurface much later, potentially contributing to the persistent faculty gender disparity (32, 33).

Current Study

In addition to determining whether faculty expressed a bias against female students, we also sought to identify the processes contributing to this bias. To do so, we investigated whether faculty members' perceptions of student competence would help to explain why they would be less likely to hire a female (relative to an identical male) student for a laboratory manager position. Additionally, we examined the role of faculty members' preexisting subtle bias against women. We reasoned that pervasive cultural messages regarding women's lack of competence in science could lead faculty members to hold gender-biased attitudes that might subtly affect their support for female (but not male) science students. These generalized, subtly biased attitudes toward women could impel faculty to judge equivalent students differently as a function of their gender.

The present study sought to test for differences in faculty perceptions and treatment of equally qualified men and women pursuing careers in science and, if such a bias were discovered, reveal its mechanisms and consequences within academic science. We focused on hiring for a laboratory manager position as the primary dependent variable of interest because it functions as a professional launching pad for subsequent opportunities. As secondary measures, which are related to hiring, we assessed: (*i*) perceived student competence; (*ii*) salary offers, which reflect the extent to which a student is valued for these competitive positions; and (*iii*) the extent to which the student was viewed as deserving of faculty mentoring.

Our hypotheses were that: Science faculty's perceptions and treatment of students would reveal a gender bias favoring male students in perceptions of competence and hireability, salary conferral, and willingness to mentor (hypothesis A); Faculty gender would not influence this gender bias (hypothesis B); Hiring discrimination against the female student would be mediated (i.e., explained) by faculty perceptions that a female student is less competent than an identical male student (hypothesis C); and Participants' preexisting subtle bias against women would moderate (i.e., impact) results, such that subtle bias against women would be negatively related to evaluations of the female student, but unrelated to evaluations of the male student (hypothesis D).

Results

A broad, nationwide sample of biology, chemistry, and physics professors (n = 127) evaluated the application materials of an undergraduate science student who had ostensibly applied for a science laboratory manager position. All participants received the same materials, which were randomly assigned either the name of a male (n = 63) or a female (n = 64) student; student gender was thus the only variable that differed between conditions. Using previously validated scales, participants rated the student's competence and hireability, as well as the amount of salary and amount of mentoring they would offer the student. Faculty participants believed that their feedback would be shared with the student they had rated (see *Materials and Methods* for details).

Student Gender Differences. The competence, hireability, salary conferral, and mentoring scales were each submitted to a two (student gender; male, female) × two (faculty gender; male, female) between-subjects ANOVA. In each case, the effect of student gender was significant (all $P < 0.01$), whereas the effect of faculty participant gender and their interaction was not (all $P > 0.19$). Tests of simple effects (all $d > 0.60$) indicated that faculty participants viewed the female student as less competent [$t(125) = 3.89$, $P < 0.001$] and less hireable [$t(125) = 4.22$, $P < 0.001$] than the identical male student (Fig. 1 and Table 1). Faculty participants also offered less career mentoring to the female student than to the male student [$t(125) = 3.77$, $P < 0.001$]. The mean starting salary offered the female student, $26,507.94, was significantly lower than that of $30,238.10 to the male student [$t(124) = 3.42$, $P < 0.01$] (Fig. 2). These results support hypothesis A.

In support of hypothesis B, faculty gender did not affect bias (Table 1). Tests of simple effects (all $d < 0.33$) indicated that female faculty participants did not rate the female student as more competent [$t(62) = 0.06$, $P = 0.95$] or hireable [$t(62) = 0.41$, $P = 0.69$] than did male faculty. Female faculty also did not offer more mentoring [$t(62) = 0.29$, $P = 0.77$] or a higher salary [$t(61) = 1.14$, $P = 0.26$] to the female student than did their male

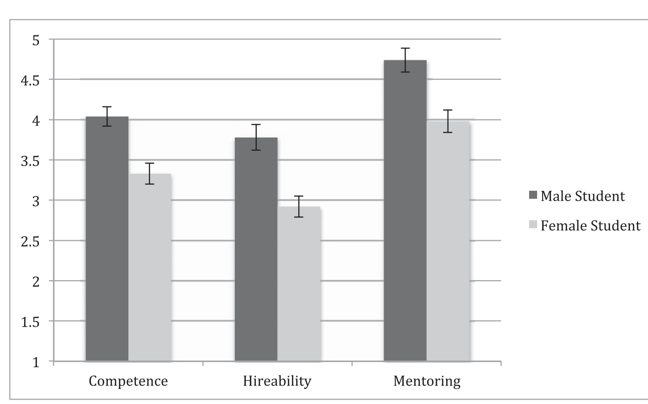

Fig. 1. Competence, hireability, and mentoring by student gender condition (collapsed across faculty gender). All student gender differences are significant ($P < 0.001$). Scales range from 1 to 7, with higher numbers reflecting a greater extent of each variable. Error bars represent SEs. $n_{\text{male student condition}} = 63$, $n_{\text{female student condition}} = 64$.

Table 1. Means for student competence, hireability, mentoring and salary conferral by student gender condition and faculty gender

	Male target student				Female target student				
	Male faculty		Female faculty		Male faculty		Female faculty		
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*d*
Competence	4.01$_a$	(0.92)	4.1$_a$	(1.19)	3.33$_b$	(1.07)	3.32$_b$	(1.10)	0.71
Hireability	3.74$_a$	(1.24)	3.92$_a$	(1.27)	2.96$_b$	(1.13)	2.84$_b$	(0.84)	0.75
Mentoring	4.74$_a$	(1.11)	4.73$_a$	(1.31)	4.00$_b$	(1.21)	3.91$_b$	(0.91)	0.67
Salary	30,520.83$_a$	(5,764.86)	29,333.33$_a$	(4,952.15)	27,111,11$_b$	(6,948.58)	25,000.00$_b$	(7,965.56)	0.60

Scales for competence, hireability, and mentoring range from 1 to 7, with higher numbers reflecting a greater extent of each variable. The scale for salary conferral ranges from \$15,000 to \$50,000. Means with different subscripts within each row differ significantly ($P < 0.05$). Effect sizes (Cohen's *d*) represent target student gender differences (no faculty gender differences were significant, all $P > 0.14$). Positive effect sizes favor male students. Conventional small, medium, and large effect sizes for *d* are 0.20, 0.50, and 0.80, respectively (51). $n_{\text{male student condition}} = 63$, $n_{\text{female student condition}} = 64$. ***$P < 0.001$.

colleagues. In addition, faculty participants' scientific field, age, and tenure status had no effect (all $P > 0.53$). Thus, the bias appears pervasive among faculty and is not limited to a certain demographic subgroup.

Mediation and Moderation Analyses. Thus far, we have considered the results for competence, hireability, salary conferral, and mentoring separately to demonstrate the converging results across these individual measures. However, composite indices of measures that converge on an underlying construct are more statistically reliable, stable, and resistant to error than are each of the individual items (e.g., refs. 34 and 35). Consistent with this logic, the established approach to measuring the broad concept of target competence typically used in this type of gender bias research is to standardize and average the competence scale items and the salary conferral variable to create one composite competence index, and to use this stable convergent measure for all analyses (e.g., refs. 36 and 37). Because this approach obscures mean salary differences between targets, we chose to present salary as a distinct dependent variable up to this point, to enable a direct test of the potential discrepancy in salary offered to the male and female student targets. However, to rigorously examine the processes underscoring faculty gender bias, we reverted to standard practices at this point by averaging the standardized salary variable with the competence scale items to create a robust composite competence variable ($\alpha = 0.86$). This composite competence variable was used in all subsequent mediation and moderation analyses.

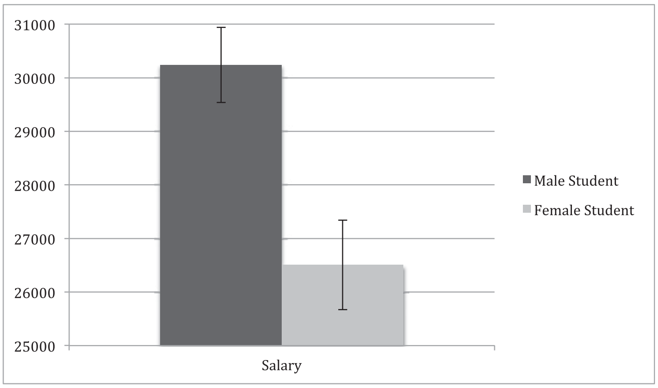

Fig. 2. Salary conferral by student gender condition (collapsed across faculty gender). The student gender difference is significant ($P < 0.01$). The scale ranges from \$15,000 to \$50,000. Error bars represent SEs. $n_{\text{male student condition}} = 63$, $n_{\text{female student condition}} = 64$.

Evidence emerged for hypothesis C, the predicted mediation (i.e., causal path; see *SI Materials and Methods: Additional Analyses* for more information on mediation and the results of additional mediation analyses). The initially significant impact of student gender on hireability ($\beta = -0.35$, $P < 0.001$) was reduced in magnitude and dropped to nonsignificance ($\beta = -0.10$, $P = 0.13$) after accounting for the impact of student composite competence (which was a strong predictor, $\beta = 0.69$, $P < 0.001$), Sobel's $Z = 3.94$, $P < 0.001$ (Fig. 3). This pattern of results provides evidence for full mediation, indicating that the female student was less likely to be hired than the identical male because she was viewed as less competent overall.

We also conducted moderation analysis (i.e., testing for factors that could amplify or attenuate the demonstrated effect) to determine the impact of faculty participants' preexisting subtle bias against women on faculty participants' perceptions and treatment of male and female science students (see *SI Materials and Methods: Additional Analyses* for more information on and the results of additional moderation analyses). For this purpose, we administered the Modern Sexism Scale (38), a well-validated instrument frequently used for this purpose (*SI Materials and Methods*). Consistent with our intentions, this scale measures unintentional negativity toward women, as contrasted with a more blatant form of conscious hostility toward women.

Results of multiple regression analyses indicated that participants' preexisting subtle bias against women significantly interacted with student gender to predict perceptions of student composite competence ($\beta = -0.39$, $P < 0.01$), hireability ($\beta = -0.31$, $P < 0.05$), and mentoring ($\beta = -0.55$, $P < 0.001$). To interpret these significant interactions, we examined the simple effects separately by student gender. Results revealed that the more preexisting subtle bias participants exhibited against women, the less composite competence ($\beta = -0.36$, $P < 0.01$) and hireability ($\beta = -0.39$, $P < 0.01$) they perceived in the female student, and the less mentoring ($\beta = -0.53$, $P < 0.001$) they were willing to offer her. In contrast, faculty participants' levels of preexisting subtle bias against women were unrelated to the perceptions of the male student's composite competence ($\beta = 0.16$, $P = 0.22$) and hireability ($\beta = 0.07$, $P = 0.59$), and the amount of mentoring ($\beta = 0.22$, $P = 0.09$) they were willing to offer him. [Although this effect is marginally significant, its direction suggests that faculty participants' preexisting subtle bias against women may actually have made them more inclined to mentor the male student relative to the female student (although this effect should be interpreted with caution because of its marginal significance).] Thus, it appears that faculty participants' preexisting subtle gender bias undermined support for the female student but was unrelated to perceptions and treatment of the male student. These findings support hypothesis D.

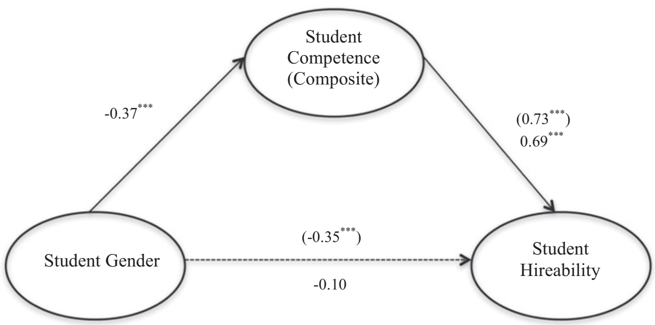

Fig. 3. Student gender difference hiring mediation. Values are standardized regression coefficients. The value in parentheses reflects a bivariate analysis. The dashed line represents the mediated path. The composite student competence variable consists of the averaged standardized salary variable and the competence scale items. Student gender is coded such that male = 0, female = 1. $n_{\text{male student condition}} = 63$, $n_{\text{female student condition}} = 64$. $^{***}P < 0.001$.

Finally, using a previously validated scale, we also measured how much faculty participants liked the student (see *SI Materials and Methods*). In keeping with a large body of literature (15), faculty participants reported liking the female (mean = 4.35, SD = 0.93) more than the male student [(mean = 3.91, SD = 0.1.08), $t(125) = -2.44$, $P < 0.05$]. However, consistent with this previous literature, liking the female student more than the male student did not translate into positive perceptions of her composite competence or material outcomes in the form of a job offer, an equitable salary, or valuable career mentoring. Moreover, only composite competence (and not likeability) helped to explain why the female student was less likely to be hired; in mediation analyses, student gender condition (β = −0.48, $P < 0.001$) remained a strong predictor of hireability along with likeability (β = 0.60, $P < 0.001$). These findings underscore the point that faculty participants did not exhibit outright hostility or dislike toward female students, but were instead affected by pervasive gender stereotypes, unintentionally downgrading the competence, hireability, salary, and mentoring of a female student compared with an identical male.

Discussion

The present study is unique in investigating subtle gender bias on the part of faculty in the biological and physical sciences. It therefore informs the debate on possible causes of the gender disparity in academic science by providing unique experimental evidence that science faculty of both genders exhibit bias against female undergraduates. As a controlled experiment, it fills a critical gap in the existing literature, which consisted only of experiments in other domains (with undergraduate students as participants) and correlational data that could not conclusively rule out the influence of other variables.

Our results revealed that both male and female faculty judged a female student to be less competent and less worthy of being hired than an identical male student, and also offered her a smaller starting salary and less career mentoring. Although the differences in ratings may be perceived as modest, the effect sizes were all moderate to large (d = 0.60–0.75). Thus, the current results suggest that subtle gender bias is important to address because it could translate into large real-world disadvantages in the judgment and treatment of female science students (39). Moreover, our mediation findings shed light on the processes responsible for this bias, suggesting that the female student was less likely to be hired than the male student because she was perceived as less competent. Additionally, moderation results indicated that faculty participants' preexisting subtle bias against women undermined their perceptions and treatment of the female (but not the male) student, further suggesting that chronic subtle biases may harm women within academic science. Use of a randomized controlled design and established practices from audit study methodology support the ecological validity and educational implications of our findings (*SI Materials and Methods*).

It is noteworthy that female faculty members were just as likely as their male colleagues to favor the male student. The fact that faculty members' bias was independent of their gender, scientific discipline, age, and tenure status suggests that it is likely unintentional, generated from widespread cultural stereotypes rather than a conscious intention to harm women (17). Additionally, the fact that faculty participants reported liking the female more than the male student further underscores the point that our results likely do not reflect faculty members' overt hostility toward women. Instead, despite expressing warmth toward emerging female scientists, faculty members of both genders appear to be affected by enduring cultural stereotypes about women's lack of science competence that translate into biases in student evaluation and mentoring.

Our careful selection of expert participants revealed gender discrimination among existing science faculty members who interact with students on a regular basis (*SI Materials and Methods: Subjects and Recruitment Strategy*). This method allowed for a high degree of ecological validity and generalizability relative to an approach using nonexpert participants, such as other undergraduates or lay people unfamiliar with laboratory manager job requirements and academic science mentoring (i.e., the participants in much psychological research on gender discrimination). The results presented here reinforce those of Stenpries, Anders, and Ritzke (40), the only other experiment we know of that recruited faculty participants. Because this previous experiment also indicated bias within academic science, its results raised serious concerns about the potential for faculty bias within the biological and physical sciences, casting further doubt on assertions (based on correlational data) that such biases do not exist (9–11). In the Steinpreis et al. experiment, psychologists were more likely to hire a psychology faculty job applicant when the applicant's curriculum vitae was assigned a male (rather than female) name (40). This previous work invited a study that would extend the finding to faculty in the biological and physical sciences and to reactions to undergraduates, whose competence was not already fairly established by accomplishments associated with the advanced career status of the faculty target group of the previous study. By providing this unique investigation of faculty bias against female students in biological and physical sciences, the present study extends past work to a critical early career stage, and to fields where women's underrepresentation remains stark (2–4).

Indeed, our findings raise concerns about the extent to which negative predoctoral experiences may shape women's subsequent decisions about persistence and career specialization. Following conventions established in classic experimental studies to create enough ambiguity to leave room for potentially biased responses (20, 23), the student applicants in the present research were described as qualified to succeed in academic science (i.e., having coauthored a publication after obtaining 2 y of research experience), but not irrefutably excellent. As such, they represented a majority of aspiring scientists, and were precisely the type of students most affected by faculty judgments and mentoring (see *SI Materials and Methods* for more discussion). Our results raise the possibility that not only do such women encounter biased judgments of their competence and hireability, but also receive less faculty encouragement and financial rewards than identical male counterparts. Because most students depend on feedback from their environments to calibrate their own worth (41), faculty's assessments of students' competence likely contribute to students' self-efficacy and goal setting as scientists, which may influence decisions much later in their careers. Likewise, inasmuch as the advice and mentoring that students receive affect their ambitions and choices, it is significant that the faculty in this study were less inclined to mentor women than men. This finding raises the possibility that women may opt out of academic science careers in part because of diminished competence judgments, rewards, and mentoring received in the early years of the careers. In sum, the predoctoral years represent a window during which students' experiences of faculty bias or encouragement are particularly likely to shape their persistence in academic science (30–33). Thus, the present study not only fills an important gap in the research literature, but also has critical implications for pressing social and educational issues associated with the gender disparity in science.

If women's decisions to leave science fields when or before they reach the faculty level are influenced by unequal treatment by undergraduate advisors, then existing efforts to create more flexible work settings (42) or increase women's identification with science (27) may not fully alleviate a critical underlying problem. Our results suggest that academic policies and mentoring interventions targeting undergraduate advisors could contribute to reducing the gender disparity. Future research should evaluate the efficacy of educating faculty and students about the existence and impact of bias within academia, an approach that has reduced racial bias among students (43). Educational efforts might address research on factors that attenuate gender bias in real-world settings, such as increasing women's self-monitoring (44). Our results also point to the importance of establishing objective, transparent student evaluation and admissions criteria to guard against observers' tendency to unintentionally use different standards when assessing women relative to men (45, 46). Without such actions, faculty bias against female undergraduates may continue to undermine meritocratic advancement, to the detriment of research and education.

Conclusions

The dearth of women within academic science reflects a significant wasted opportunity to benefit from the capabilities of our best potential scientists, whether male or female. Although women have begun to enter some science fields in greater numbers (5), their mere increased presence is not evidence of the absence of bias. Rather, some women may persist in academic science despite the damaging effects of unintended gender bias on the part of faculty. Similarly, it is not yet possible to conclude that the preferences for other fields and lifestyle choices (9–11) that lead many women to leave academic science (even after obtaining advanced degrees) are not themselves influenced by experiences of bias, at least to some degree. To the extent that faculty gender bias impedes women's full participation in science, it may undercut not only academic meritocracy, but also the expansion of the scientific workforce needed for the next decade's advancement of national competitiveness (1).

Materials and Methods

Participants. We recruited faculty participants from Biology, Chemistry, and Physics departments at three public and three private large, geographically diverse research-intensive universities in the United States, strategically selected for their representative characteristics (see *SI Materials and Methods* for more information on department selection). The demographics of the 127 respondents corresponded to both the averages for the selected departments and faculty at all United States research-intensive institutions, meeting the criteria for generalizability even from nonrandom samples (see *SI Materials and Methods* for more information on recruitment strategy and participant characteristics). Indeed, we were particularly careful to obtain a sample representative of the underlying population, because many past studies have demonstrated that when this is the case, respondents and nonrespondents typically do not differ on demographic characteristics and responses to focal variables (47).

Additionally, in keeping with recommended practices, we conducted an a priori power analysis before beginning data collection to determine the optimal sample size needed to detect effects without biasing results toward obtaining significance (*SI Materials and Methods: Subjects and Recruitment Strategy*) (48). Thus, although our sample size may appear small to some readers, it is important to note that we obtained the necessary power and representativeness to generalize from our results while purposefully avoiding an unnecessarily large sample that could have biased our results toward a false-positive type I error (48).

Procedure. Participants were asked to provide feedback on the materials of an undergraduate science student who stated their intention to go on to graduate school, and who had recently applied for a science laboratory manager position. Of importance, participants believed they were evaluating a real student who would subsequently receive the faculty participants' ratings as feedback to help their career development (see *SI Materials and Methods* for more information, and Fig. S1 for the full text of the cover story). Thus, the faculty participants' ratings were associated with definite consequences.

Following established practices, the laboratory manager application was designed to reflect high but slightly ambiguous competence, allowing for variability in participant responses (20, 23). In addition, a promising but still-nascent applicant is precisely the type of student whose persistence in academic science is most likely to be affected by faculty support or discouragement (30–33), rendering faculty reactions to such a student of particular interest for the present purposes. The materials were developed in consultation with a panel of academic science researchers (who had extensive experience hiring and supervising student research assistants) to ensure that they would be perceived as realistic (*SI Materials and Methods*). Results of a funneled debriefing (49) indicated that this was successful; no participant reported suspicions that the target was not an actual student who would receive their evaluation.

Participants were randomly assigned to one of two student gender conditions: application materials were attributed to either a male student (John, $n = 63$), or a female student (Jennifer, $n = 64$), two names that have been pretested as equivalent in likability and recognizeability (50). Thus, each participant saw only one set of materials, from either the male or female applicant (see Fig. S2 for the full text of the laboratory manager application and *SI Method and Materials* for more information on all materials). Because all other information was held constant between conditions, any differences in participants' responses are attributable to the gender of the student. Using validated scales, participants rated student competence, their own likelihood of hiring the student, selected an annual starting salary for the student, indicated how much career mentoring they would provide to such a student, and completed the Modern Sexism Scale.

ACKNOWLEDGMENTS. We thank faculty members from six anonymous universities for their involvement as participants; and Jessamina Blum, John Crosnick, Jennifer Frederick, Jaime Napier, Jojanneke van der Toorn, Tiffany Tsang, Tessa West, James Young, and two anonymous reviewers for valuable input. This research was supported by a grant from the Howard Hughes Medical Institute Professors Program (to J.H.).

1. President's Council of Advisors on Science and Technology (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Available at http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_feb.pdf. (Accessed February 13, 2012).
2. Handelsman J, et al. (2005) Careers in science. More women in science. *Science* 309:1190–1191.
3. United States National Academy of Sciences (2007) *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering* (National Academies, Washington, DC).
4. National Science Foundation (2009) *Women, Minorities, and Persons with Disabilities in Science and Engineering* (National Science Foundation, Arlington).
5. Bell N (2010) *Graduate Enrollment and Degrees: 1999 to 2009* (Council of Graduate Schools, Washington, DC).
6. Halpern DF, et al. (2007) The science of gender differences in science and mathematics. *Psychol Sci* 8(1):1–51.
7. Hyde JS, Linn MC (2006) Diversity. Gender similarities in mathematics and science. *Science* 314:599–600.
8. Spelke ES (2005) Sex differences in intrinsic aptitude for mathematics and science?: A critical review. *Am Psychol* 60:950–958.
9. Ceci SJ, Williams WM (2010) Gender differences in math-intensive fields. *Curr Dir Psychol Sci* 19:275–279.
10. Ceci SJ, Williams WM (2011) Understanding current causes of women's underrepresentation in science. *Proc Natl Acad Sci USA* 108:3157–3162.

11. Ceci SJ, et al. (2011) Do subtle cues about belongingness constrain women's career choices? *Psychol Inq* 22:255–258.
12. Berezow AB (2011) *Gender discrimination in science is a myth.* Available at http://www.nationalreview.com/articles/256816/gender-discrimination-science-myth-alex-b-berezow?pg=2. (Accessed June 6, 2012).
13. Dickey Zakaib G (2011) Science gender gap probed. *Nature* 470:153.
14. Devine PG (1989) Stereotypes and prejudice: Their automatic and controlled components. *J Pers Soc Psychol* 56(1):5–18.
15. Eagly AH, Mladinic A (1994) Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *Eur Rev Soc Psychol* 5(1):1–35.
16. Spence JT, Hahn ED (1997) The attitudes toward women scale and attitude change in college students. *Psychol Women Q* 21(1):17–34.
17. Dovidio JF, Gaertner SL (2004) *Advances in Experimental Social Psychology*, ed Zanna MP (Elsevier, New York), pp 1–51.
18. Nosek BA, Banaji M, Greenwald AG (2002) Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dyn* 6(1):101–115.
19. Goldin C, Rouse C (2000) Orchestrating impartiality: The impact of "blind" auditions on female musicians. *Am Econ Rev* 90:715–741.
20. Foschi M (2000) Double standards for competence: Theory and research. *Annu Rev Sociol* 26(1):21–42.
21. Foschi M (1996) Double standards in the evaluation of men and women. *Soc Psychol Q* 59:237–254.
22. Goldberg P (1968) Are women prejudiced against women? *Transaction* 5(5):28–30.
23. Heilman ME, Wallen AS, Fuchs D, Tamkins MM (2004) Penalties for success: Reactions to women who succeed at male gender-typed tasks. *J Appl Psychol* 89:416–427.
24. Monin B, Miller DT (2001) Moral credentials and the expression of prejudice. *J Pers Soc Psychol* 81:33–43.
25. Uhlmann EL, Cohen GL (2007) "I think it, therefore it's true": Effects of self-perceived objectivity on hiring discrimination. *Organ Behav Hum Dec* 104:207–223.
26. Nosek BA, Banaji MR, Greenwald AG (2002) Math = male, me = female, therefore math not = me. *J Pers Soc Psychol* 83:44–59.
27. Dasgupta N (2011) Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model. *Psychol Inq* 22:231–246.
28. The Massachusetts Institute of Technology (1999) *A study on the status of women faculty in science at MIT.* Available at http://web.mit.edu/fnl/women/women.html#The%20Study (Accessed February 13, 2012).
29. Steele J, James JB, Barnett RC (2002) Learning in a man's world: Examining the perceptions of undergraduate women in male-dominated academic areas. *Psychol Women Q* 26(1):46–50.
30. Lent RW, et al. (2001) The role of contextual supports and barriers in the choice of math/science educational options: A test of social cognitive hypotheses. *J Couns Psychol* 48:474–483.
31. Seymour E, Hewitt NM (1996) *Talking about Leaving: Why Undergraduates Leave the Sciences* (Westview, Boulder).
32. Gasiewski JA, et al. (2012) From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Res Higher Educ* 53:229–261.
33. Byars-Winston A, Gutierrez B, Topp S, Carnes M (2011) Integrating theory and practice to increase scientific workforce diversity: A framework for career development in graduate research training. *CBE Life Sci Educ* 10:357–367.
34. Campbell DT, Fiske DW (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 56:81–105.
35. Robins RW, Hendin HM, Trzesniewski KH (2001) Measuring global self esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Pers Soc Psychol Bull* 27(2):151–161.
36. Moss-Racusin CA, Rudman LA (2010) Disruptions in women's self-promotion: The backlash avoidance model. *Psychol Women Q* 34:186–202.
37. Rudman LA, Moss-Racusin CA, Glick P, Phelan JE (2012) *Advances in Experimental Social Psychology*, eds Devine P, Plant A (Elsevier, New York), pp 167–227.
38. Swim JK, Aikin KJ, Hall WS, Hunter BA (1995) Sexism and racism: Old-fashioned and modern prejudices. *J Pers Soc Psychol* 68:199–214.
39. Martell RF, Lane DM, Emrich C (1996) Male-female differences: A computer simulation. *Am Psychol* 51(2):157–158.
40. Steinpreis RE, Anders KA, Ritzke D (1999) The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles* 41:509–528.
41. Bandura A (1982) Self-efficacy mechanism in human agency. *Am Psychol* 37:122–147.
42. Scandura TA, Lankau MJ (1997) Relationships of gender, family responsibility and flexible work hours to organizational commitment and job satisfaction. *J Organ Behav* 18:377–391.
43. Rudman LA, Ashmore RD, Gary ML (2001) "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *J Pers Soc Psychol* 81:856–868.
44. O'Neill OA, O'Reilly CA (2011) Reducing the backlash effect: Self-monitoring and women's promotions. *J Occup Organ Psychol* 84:825–832.
45. Uhlmann EL, Cohen GL (2005) Constructed criteria: Redefining merit to justify discrimination. *Psychol Sci* 16:474–480.
46. Phelan JE, Moss-Racusin CA, Rudman LA (2008) Competent yet out in the cold: Shifting criteria for hiring reflects backlash toward agentic women. *Psychol Women Q* 32:406–413.
47. Holbrook AL, Krosnick JA, Pfent A (2007) *Advances in Telephone Survey Methodology*, eds Lepkowski JM, et al. (John Wiley & Sons, Hoboken), pp 499–528.
48. Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 22:1359–1366.
49. Bargh JA, Chartrand TL (2000) *Handbook of Research Methods in Social and Personality Psychology*, eds Reis HT, Judd CM (Cambridge Univ Press, New York), pp 253–285.
50. Brescoll VL, Uhlmann EL (2008) Can angry women get ahead? Status conferral, gender, and workplace emotion expression. *Psychol Sci* 19:268–275.
51. Cohen J (1998) *Statistical Power Analysis for the Behavioral Sciences* (Erlbaum, Hillsdale), 2nd Ed.

Supporting Information

Moss-Racusin et al. 10.1073/pnas.1211286109

SI Materials and Methods

Subjects and Recruitment Strategy. To identify and screen potential participants, we used established practices similar to those used in other field experiments relying on nonundergraduate samples (1). We sought to strategically select departments for inclusion that were representative of high-quality United States science programs. Thus, participants were recruited from six anonymous American universities, all of which were ranked by the Carnegie Foundation as "large, Research University (very high research productivity)" (2). Additionally, each university had prominent, well-respected science departments (both at the undergraduate and graduate level), and tended to graduate high numbers of students who go on to careers in academic science. The schools were matched for size and prestige, and were selected from three different geographic regions within the United States. Within each region we included one private and one public university.

Within each university, participants were recruited from Biology, Chemistry, and Physics departments. These three fields were chosen because of their size, prominence, competitiveness, emphasis on research, and varying gender disparities. That is, all although all three showed gender disparities at the faculty level, the size of the gap differed. This gap was even more pronounced at the doctoral level, with some subfields of the biological sciences granting more doctorates to women than men (3). This diversity allowed for an examination of how faculty bias might differ as a function of the size and severity of the gender disparity across science fields. Although each institution had only one clear Chemistry and Physics department, some institutions had more than one Biology department (e.g., Ecology and Evolutionary Biology; Molecular, Cellular and Developmental Biology, and so forth). For such institutions, each of the core Biology departments were included. This method yielded a total of 23 departments.

For each selected department, departmental Web sites and publicly available course listings were used to create a full list of all eligible faculty participants. Potential participants had to meet several a priori qualifications to be eligible for participation. First, the participants had to be clearly identified as tenure-track faculty. As a result, Visiting Assistant Professors, Adjunct Professors, Instructional Staff, Research Faculty, Postdoctoral Associates, and Lecturers were not included. Additionally, participants had to be current faculty, thus excluding those with Emeritus titles. Faculty with primary appointments in other departments were also excluded, as were those whose appointments had yet to officially begin and those with invalid e-mail addresses. Finally, faculty who were identified as close personal friends or colleagues of one of the present study's authors were eliminated to avoid conflicts of interest. This method yielded a total of 547 eligible participants.

Data were collected during October and November of 2011. We followed the general methodological approach used in correspondence test audit research, typically used in field studies of discrimination (4, 5). All eligible participants received an e-mailed participation invitation originating from C.A.M.-R., indicated their consent, and completed all measures online. This method yielded a total of 165 participants, for an overall response rate of 30% (percentage rounded up). This percentage is on par with both similar empirical studies of professionals (6, 7) and that typically obtained in survey research (8). Additionally, extensive previous research has indicated that both demographic characteristics and substantive responses to focal variables largely do not differ between respondents and nonrespondents when sample demographics correspond to those of the underlying population (8). Thus, because the demographic information of our participants reflected the underlying population (as discussed below), the response rate obtained in the present study should allow for reasonably generalizable conclusions.

Data obtained from 30 participants were used to pilot and improve the study instruments and were thus not included in final analyses. Of the remaining 135 participants, 8 did not complete the majority of the study because of computer error (in three cases) or attrition (five cases); this resulted in a final sample of 127 participants for all substantive analyses. A power analysis indicated that this sample size exceeded the recommended $n = 90$ required to detect moderate effect sizes. Of participants, 74% were male and 81% were White (specific ethnic backgrounds were reported as follows: 81% White, 6% East-Asian, 4% South-Asian, 2% Hispanic, 2% African-American, 2% multiracial, and 1% each for Southeast-Asian, Middle-Eastern, and other), with a mean age of 50.34 (SD = 12.60, range 29–78). Of importance, these demographics are representative of both the averages for the 23 sampled departments (demographic characteristics for the sampled departments were 78% male and 81% White, corresponding closely with the demographics of those who elected to participate), as well as national averages (9). Additionally, 18% of participants were Assistant Professors, 22% were Associate Professors, and 60% were full professors, with 40% Biologists, 32% Physicists, and 28% Chemists. No demographic variables were associated with participants' substantive responses (all $P >$ 0.53). As expected when using random assignment, participants' demographic conditions did not vary across experimental conditions. Because there were 15 female and 48 male participants in the male student condition, and 18 female and 45 male participants in the female student condition, we obtained sufficient power to test our hypotheses (10).

Student Laboratory Manager Application Materials. We asked participants to rate a student laboratory manager application to help us develop appropriate mentoring programs for undergraduate science students. We prefaced the student laboratory manager application with text designed to bolster the credibility of the cover story and adjust for any differences in expectations and practices regarding laboratory managers between science fields (Fig. S1). Following conventions established in previous experimental work (11, 12), the laboratory manager application was designed to reflect slightly ambiguous competence, allowing for variability in participant responses and the utilization of biased evaluation strategies (if they exist). That is, if the applicant had been described as irrefutably excellent, most participants would likely rank him or her highly, obscuring the variability in responses to most students for whom undeniable competence is frequently not evident. Even if gender-biased judgments do typically exist when faculty evaluate most undergraduates, an extraordinary applicant may avoid such biases by virtue of their record. This approach also maintained the ecological validity and generalizability of results to actual undergraduate students of mixed ability levels.

Thus, we followed procedures established in previous similar research (11, 12) by designing an applicant who was "in the ballpark" for a laboratory manager position, but was not an obvious star. For example, although the applicant had completed 2 y of research experience and coauthored a journal article, their grade point average was slightly low (3.2) and they were described as having withdrawn from one class before the final. Fig. S2 displays the full text of the student laboratory manager application materials for the female student condition. The sole difference in the male student condition was that the student's name read as "John" instead of "Jennifer," and female pronouns were replaced with male pronouns.

To ensure that the application materials reflected the desired degree of competence, they were developed in consultation with a panel of academic science researchers who had extensive experience hiring and supervising student research assistants. After the materials were developed, they were then rated by a separate group of knowledgeable graduate students, postdoctoral scholars, and faculty. Results from this pilot testing revealed consensus that, as intended, the materials reflected a qualified but not irrefutably excellent applicant.

Dependent Variable Scales. Participants completed the following scales, which were well-validated and modified for use from previous studies (13–15).

Student competence. The target student's competence was assessed using three items on a 1 (not at all) to 7 (very much) scale. These items were: (*i*) Did the applicant strike you as competent? (*ii*) How likely is it that the applicant has the necessary skills for this job? (*iii*) How qualified do you think the applicant is? ($\alpha = 0.93$). Items were averaged to form the student competence scale, with higher numbers indicating greater levels of perceived competence.

Student hireability. The extent to which the student applicant was viewed as hireable for a laboratory manager position was measured using three items on a 1 (not at all likely) to 7 (very likely) scale. These items were: (*i*) How likely would you be to invite the applicant to interview for the laboratory manager job? (*ii*) How likely would you be to hire the applicant for the laboratory manager job? (*iii*) How likely do you think it is that the applicant was actually hired for the laboratory manager job he/she applied for? ($\alpha = 0.91$). Items were averaged to compute the student hireability scale, such that higher numbers reflected greater perceived hireability.

Salary conferral. Salary conferral was measured using one item, If you had to choose one of the following starting salaries for the applicant, what would it be? Responses were indicated on the following scale: 1 ($15,000), 2 ($20,000), 3 ($25,000), 4 ($30,000), 5 ($35,000), 6 ($40,000), 7 ($45,000), 8 ($50,000). Collapsed across conditions, the average recommended salary was $28,373.02 (SD = $6,382.14), with a range of $15,000 to $45,000.

Mentoring. The extent to which participants were willing to mentor the student applicant was assessed using three items on a 1 (not at all likely) to 7 (very likely) scale. These items were: If you encountered this student at your own institution, how likely would you be to… (*i*) Encourage the applicant to stay in the field if he/she was considering changing majors? (*ii*) Encourage the applicant to continue to focus on research if he/she was considering switching focus to teaching? (*iii*) Give the applicant extra help if he/she was having trouble mastering a difficult concept? ($\alpha = 0.73$). Items were averaged to form the mentoring scale, with high numbers reflecting greater willingness to mentor the student.

Subtle gender bias. Our intention was to select a scale that would measure modern bias against women. We reasoned that, as mentors and educators, many faculty members would not report high levels of "old-fashioned" or hostile sexism, characterized by overtly negative evaluations of women and the desire to halt women's progress (16). Instead, we were interested in how shared subtle gender biases (resulting from pervasive cultural messages) might impact perceptions of female students. As a result, we sought a measure that would tap a subtle, modern form of gender bias that often exists outside of individuals' conscious awareness or intention to harm women. Thus, we used the Modern Sexism Scale, a commonly used and well-validated scale that functions as an indirect measure of modern views toward women and gender (17). Participants responded to eight items on a scale ranging from 1 (strongly disagree) to 7 (strongly agree). Items included: On average, people in our society treat husbands and wives equally; Discrimination against women is no longer a problem in the United States; and Over the past few years, the government and new media have been showing more concern about the treatment of women than is warranted by women's actual experiences ($\alpha = 0.92$). Items were averaged to form the gender attitudes scale, with higher numbers indicating more negative attitudes toward women.

Likeability. Using a scale ranging from 1 (not at all likely) to 7 (very likely) scale, participants answered three items indicating the extent to which they liked the student applicant. These were: (*i*) How much did you like the applicant? (*ii*) Would you characterize the applicant as someone you want to get to know better? (*iii*) Would the applicant fit in well with other laboratory members? Items were averaged to create the likeability scale, with higher numbers representing greater liking of the target student ($\alpha = 0.87$).

Analytic Strategy. Although we treated the individual as the primary unit of analysis for all statistical tests, it should be noted that our data may be conceptualized as clustered or nested by various groups (18). That is, because participants belonged to one of three distinct Science fields (Biology, Chemistry, or Physics) and also one of 23 distinct departments, their membership in a science field or a department could result in nonindependence of observations on at least one of these levels. For example, if bias against female undergraduates is systematically greater among chemists than biologists, then our data would be nonindependent at the level of science field (in that scores on any given dependent variable would be more similar for two chemists than two randomly selected individuals) (18). Because standard inferential statistics assume that all observations are independent, we may have introduced error by failing to account for the nested nature of our data (i.e., nonindependence because of groups) (19).

To address this issue, we followed recommended practices to assess the possible nonindependence of our data (20). If data are found to be independent at the level of a nesting variable, then it is acceptable to not account for this variable in statistical tests (18, 19). To evaluate the nonindependence of our data, we conducted a null multilevel model for each dependent variable. Using the MIXED procedure in SPSS, we included both predictors (student and faculty gender, effects-coded such that male = −1, female = 1) in each model. For each model, the intraclass correlation coefficient was near zero and nonsignificant (all $P > 0.11$), suggesting that our data did not violate assumptions of independence. Put another way, there was no significant variance associated with participants' group membership in a given science field. As a result, we concluded that it was appropriate to analyze our data without accounting for the impact of this nesting variable.

One additional potential source of group level variance remains. Our data were also nested at the level of department. Although it is possible that participants' scores were nonindependent based on departmental membership, we were not able to estimate nonindependence at this level because of ethical concerns that shaped our recruitment and data-collection strategies. That is, following the stipulation of the reviewing Institutional Review Board, we assured full anonymity to our faculty participants. As a result, it was impossible to ask them to indicate to which department they belonged. We agreed with the Institutional Review Board's assessment that participation and attrition rates as well as responses to potentially sensitive questions may have been biased if participants feared that their identities could be gleaned from the information they provided. Indeed, among (sometimes small) academic fields, knowing an individual's university, department, and demographic characteristics would likely be sufficient information to identify them personally.

Thus, to avoid undermining recruitment strategies and biasing participants' responses, we refrained from collecting information about specific departmental membership. As such, we were unable to determine whether responses were nonindependent as a function of this variable, and if so, to account for it in our statistical approach. Ignoring significant nonindependence does not bias effect estimates, but can bias degrees of freedom, variances, SEs, and significance tests (20). Monte Carlo simulations have determined that under these circumstances, the resulting SEs may be too large, too small, or hardly biased, resulting in possible type 1, type II, or no errors (21). However, because the data were shown to be nonindependent at the level of science field and there is no a priori theoretical reason to predict nonindependence at the level of department, we proceeded with a standard inferential statistical approach. Nonetheless, future research should seek to develop methods to measure and control for potential nonindependence due to departmental membership.

Additional Analyses. ***Mediation analyses.*** To test for mediation, we followed procedures recommended by Baron and Kenny (22). Student gender was coded 0 (male), 1 (female). As noted in the main text, primary mediation analyses evaluating hypothesis C were conducted using the composite competence variable, following established best practices (e.g., refs. 23–26). However, we also ensured that results were similar using the noncomposite competence scale (i.e., without the salary conferral variable), to rule out the possibility that the mediation results were driven solely or primarily by the salary conferral variable. As expected, results using the noncomposite competence scale were similar although slightly weaker, in that the initially significant relationship between student gender and student hireability ($\beta = -0.35$, $P < 0.001$) was reduced in magnitude and significance ($\beta = -0.13$, $P = 0.05$) after accounting for the impact of student competence (which was a strong predictor, $\beta = 0.69$, $P < 0.001$), Sobel's $Z = 3.65$, $P < 0.001$. These results suggest that the composite competence variable functioned as expected.

Additionally, although not specifically predicted, we examined whether the composite competence variable might also mediate the relationship between student gender and mentoring. Results demonstrated partial mediation, in that the initially significant relationship between student gender and mentoring ($\beta = -0.32$, $P < 0.001$) was reduced in magnitude and significance ($\beta = -0.22$, $P = 0.02$) after accounting for the impact of student composite competence (which was a significant predictor, $\beta = 0.28$, $P < 0.05$), Sobel's $Z = 2.91$, $P < 0.01$.

Because mentoring has strong social components and may be perceived as less immediately task-relevant than hireability, we did not initially expect it to be mediated by the composite competence scale, and this could account for why we observed partial rather than full mediation with this variable (relative to the full mediation observed for hireability, supporting hypothesis C). However, the fact that evidence for partial mediation emerged even for mentoring (a secondary downstream dependent variable in the current context) speaks to the powerful impact of differences in the perceived competence of male and female students.

Moderation analyses. To test for moderation, we first standardized all variables and then ran a series of multiple regression analyses with student gender, faculty participants' negative attitudes toward women, and their interaction predicting student composite competence, hireability, and mentoring. As noted, the interaction was a significant predictor in each case. As with mediation, we next ensured that results were similar using the noncomposite measure of competence (to determine that moderation results for this variable were not driven solely by the salary conferral variable, which was included in the composite competence measure). Results of multiple regression analyses indicated that participants' preexisting subtle bias against women significantly interacted with student gender condition to predict perceptions of student noncomposite competence ($\beta = -0.42$, $P < 0.01$). As expected, bivariate analyses results revealed that the more preexisting subtle bias participants exhibited against women, the less noncomposite competence ($\beta = -0.38$, $P < 0.01$) they perceived the female student to possess. In contrast, faculty participants' levels of preexisting subtle bias against women were unrelated to the perceptions of the male student's noncomposite competence ($\beta = 0.18$, $P = 0.16$). These results are nearly identical to those obtained with the composite competence index, suggesting that those findings were not driven solely by the salary conferral variable and providing additional evidence that the composite competence variable functioned as intended.

We then explored whether additional variables might interact with participants' subtle preexisting bias against women to predict their reactions to the target students. Separate models adding faculty participant gender, age, science field, tenure status, each two-way interaction, as well as the three-way interaction of each demographic variable with student gender condition and faculty participant gender (to rule out participant gender differences) revealed no significant novel predictors (all $\beta < 0.38$, all $P > 0.28$). This finding suggests that faculty participants' gender attitudes themselves played a role in undermining support for the female (but not male) student, and that the impact of these gender attitudes does not appear to vary as a function of participants' other demographic characteristics, including their gender. Consistent with other results, it appears that female as well as male faculty members' negative attitudes toward women undermined their support for the female student, irrespective of their age, science field, and career status.

1. Feldon DF, et al. (2011) Graduate students' teaching experiences improve their methodological research skills. *Science* 333:1037–1039.
2. Carnegie Foundation for the Advancement of Teaching (2012) *Standard listing of research universities (very high research productivity).* Available at http://classifications.carnegiefoundation.org/lookup_listings/srp.php?clq=%7B%22basic2005_ids%22%3A%2215%22%7D&start_page=standard.php&backurl=standard.php&limit=0,50. (Accessed February 14, 2012).
3. National Science Foundation (2009) *Women, Minorities, and Persons with Disabilities in Science and Engineering* (National Science Foundation, Arlington).
4. Pager D (2007) The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *Ann Am Acad Pol Soc Sci* 609 (1):104–133.
5. Bertrand M, Mullainathan S (2004) Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *Am Econ Rev* 94: 991–1013.
6. Steinpreis RE, Anders KA, Ritzke D (1999) The impact of gender on the review of curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles* 41:509–528.
7. Wunder GC, Wynn GW (1988) The effects of address personalization on mailed questionnaires response rate, time, and quality. *J Mark Res Soc* 30(1):9–101.
8. Holbrook AL, Krosnick JA, Pfent A (2007) *Advances in Telephone Survey Methodology*, eds Lepkowski JM, et al. (John Wiley & Sons, Hoboken), pp 499–528.
9. National Science Foundation (2008) *Survey of Earned Doctorates* (National Science Foundation, Arlington).
10. Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 22:1359–1366.
11. Foschi M (2000) Double standards for competence: Theory and research. *Annu Rev Sociol* 26(1):21–42.
12. Heilman ME, Wallen AS, Fuchs D, Tamkins MM (2004) Penalties for success: Reactions to women who succeed at male gender-typed tasks. *J Appl Psychol* 89:416–427.
13. Moss-Racusin CA, Phelan JE, Rudman LA (2010) When men break the gender rules: Status incongruity and backlash against modest men. *Psychol Men Masc* 11(2):140–151.
14. Moss-Racusin CA, Rudman LA (2010) Disruptions in women's self-promotion: The backlash avoidance model. *Psychol Women Q* 34:186–202.
15. Rudman LA, Moss-Racusin CA, Phelan JE, Nauts S (2012) Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice toward female leaders. *J Exp Soc Psychol* 48(1):165–179.
16. Glick P, Fiske ST (1996) The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *J Pers Soc Psychol* 70:491–512.
17. Swim JK, Aikin KJ, Hall WS, Hunter BA (1995) Sexism and racism: Old-fashioned and modern prejudices. *J Pers Soc Psychol* 68:199–214.
18. Kenny DA (1996) The design and analysis of social-interaction research. *Annu Rev Psychol* 47:59–86.

19. Kenny DA, Judd CM (1986) Consequences of violating the independence assumption in analysis of variance. *Psychol Bull* 99:422–431.
20. Kenny DA, Judd CM (1996) A general procedure for the estimation of interdependence. *Psychol Bull* 119:138–148.
21. Hox J (1998) *Classification, Data Analysis, and Data Highways*, eds Balderjahn I, Mather R, Schader M (Springer, New York), pp 147–154.
22. Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 51:1173–1182.
23. Campbell DT, Fiske DW (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 56:81–105.
24. Robins RW, Hendin HM, Trzesniewski KH (2001) Measuring global self esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Pers Soc Psychol Bull* 27(2):151–161.
25. Moss-Racusin CA, Rudman LA (2010) Disruptions in women's self-promotion: The backlash avoidance model. *Psychol Women Q* 34:186–202.
26. Rudman LA, Moss-Racusin CA, Glick P, Phelan JE (2012) *Advances in Experimental Social Psychology*, eds Devine P, Plant A (Elsevier, New York), pp 167–227.

In the current phase of the study, we are interested in how undergraduate students' are selected for competitive lab manager positions following their graduation and before they apply to doctoral programs. Lab managers are an important part of the scientific enterprise, but very little is know about how people weigh various factors and make the complicated decision to hire one lab manager from a pool of applicants.

To study this question, we have compiled and summarized information from ***actual applications of students who have recently applied to be lab managers*** *at universities across the country. These students have volunteered to share their information in exchange for mentoring opportunities as part of their participation in the study. We have summarized their information by creating profiles for each applicant using a standardized form, in order to adjust for individual differences in application procedures and enable consistent evaluations across applicants.*

Today, we will be assigning you to read the applicant profile of one randomly-selected student from the nationwide database. Please imagine that you are actually evaluating this student's application to work ***in your own lab****. After reading the applicant profile, you will be asked to provide your opinions of the student, and offer them feedback as they make decisions about moving forward with their career. These students come from fields where lab managers are typically students with research-focused career ambitions who work in these types of positions for several years before going on to graduate school. Even if your particular sub-field does not have such practices regarding lab managers, please evaluate the student within this context.*

Remember, the national database we have created contains information both from very qualified students, as well as students who are less competitive. Thus, we ask that you be prepared to offer honest feedback to the student you are selected to read about, regardless of how qualified they are. Please do not hold back, and don't be afraid to be truthful—the success of the current project depends upon your ability to provide honest, straightforward feedback.

Fig. S1. Cover story text. The text in the figure was viewed by participants in PDF format without additional supporting text.

DEMOGRAPHICS
Participant ID #: 149
Name: Jennifer [redacted]
Gender: Female
Ethnic Background: Caucasian
Age: 22
Degree: Bachelors of Science, obtained May 2011 from [redacted] University

BACKGROUND
GPA: 3.2
GRE score: 650 verbal, 780 quant
Awards/honors: President's Service Award, Rotary Club College Scholarship
Previous research experience: 2 years as a research assistant working with 2 different faculty mentors
Academic standing: appears from Jennifer's transcript that she was in good standing upon graduation, but withdrew from 1 class prior to final
Letters of recommendation: 3 (2 from former faculty research supervisors, 1 from an intro science course professor), all supportive
Future plans: apply to doctoral programs
Extracurricular activities: student government, college learning center tutor
Position sought: Lab Manager
Position duration: 2 years, with possibility of renewal pending satisfactory performance

STATEMENTS/LETTERS
Excerpt from student statement: "I am a motivated student and would make the most of the opportunity to serve as your lab manager. After spending a semester working in Dr. [redacted]'s lab and another year doing research with Dr. [redacted], I have gained valuable technical skills, co-authored a journal article, and am now committed to an academic research career…as someone focused on improving my standing and enhancing my research experience, this lab manager position would provide the perfect opportunity to hone the necessary skills to make me competitive for graduate school applications… additionally, the fascinating research taking place in your lab is directly in line with my interests and experiences…in short, I am focused, motivated, organized and dedicated to improving my research skills. I am enthusiastic about the opportunity to fill the lab manager position and collaborate with you on future research."

Excerpt from faculty recommendation letter: "…although Jennifer admittedly took a bit longer than some students to get serious about her studies early in college, she has impressed me by improving over the last two years of her science coursework and has made every effort to make up for lost ground…she has been a strong research assistant in my lab, and I know she is capable of serving as a dedicated lab manager."

Fig. S2. Lab manager application materials (female student condition). The only differences in the male student condition were that the name "Jennifer" was replaced with "John," and all female pronouns were replaced with male pronouns.