Journal of Experimental Social Psychology 73 (2017) 211–215

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

FlashReport

A gender bias habit-breaking intervention led to increased hiring of female faculty in STEMM departments☆,☆☆

Patricia G. Devine[a,*], Patrick S. Forscher[b], William T.L. Cox[a], Anna Kaatz[c], Jennifer Sheridan[d], Molly Carnes[c]

[a] Department of Psychology, University of Wisconsin-Madison, United States
[b] Department of Psychology, University of Arkansas, United States
[c] Center for Women's Health Research, University of Wisconsin-Madison, United States
[d] Women in Science and Engineering Leadership Institute, University of Wisconsin-Madison, United States

ARTICLE INFO

Keywords:
Gender disparities in STEMM
Prejudice reduction
Diversity
STEMM
Interventions
Sexism
Hiring
Gender

ABSTRACT

Addressing the underrepresentation of women in science is a top priority for many institutions, but the majority of efforts to increase representation of women are neither evidence-based nor rigorously assessed. One exception is the gender bias habit-breaking intervention (Carnes et al., 2015), which, in a cluster-randomized trial involving all but two departmental clusters ($N = 92$) in the 6 STEMM focused schools/colleges at the University of Wisconsin–Madison, led to increases in gender bias awareness and self-efficacy to promote gender equity in academic science departments and perceptions of a more positive departmental climate. Following this initial success, the present study compares, in a preregistered analysis, hiring rates of new female faculty pre- and post-manipulation. Whereas the proportion of women hired by control departments remained stable over time, the proportion of women hired by intervention departments increased by an estimated 18 percentage points ($OR = 2.23$, $d_{OR} = 0.34$). Though the preregistered analysis did not achieve conventional levels of statistical significance ($p < 0.07$), the study has a hard upper limit on statistical power, as the cluster-randomized trial has a maximum sample size of 92 departmental clusters. These findings, however, have undeniable practical significance for the advancement of women in science, and provide promising evidence that psychological interventions can facilitate gender equity and diversity.

Women remain underrepresented in doctoral-level careers in science, technology, engineering, math, and medical (STEMM) fields (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012; NSF, 2007). This gender inequity, paired with concurrent underrepresentation of racial minorities, has led numerous organizations to call for efforts to increase participation of women and minorities in STEMM (e.g., NSF, 2014; National Academy of Sciences, National Academy of Engineering, Institute of Medicine of the National Academies, 2006; NIH: Valantine & Collins, 2015; see also Corrice, 2009; Hill, Corbett, & St. Rose, 2010; Mitchneck, Smith, & Latimer, 2016; Sevo & Chubin, 2008). Many existing efforts to address these issues, however, are neither evidence-based nor rigorously assessed in experimental trials (Moss-Racusin et al., 2014; Paluck & Green, 2009). When systematically assessed, these non-evidence-based efforts either do not work or make problems worse (Apfelbaum, Norton, & Sommers, 2012; Dobbin & Kalev, 2013; Legault, Gutsell, & Inzlicht, 2011).

Interventions designed to reduce intergroup biases should be rooted in well-supported theory about the nature of prejudice and bias reduction. One such theory is the prejudice habit model (Devine, 1989; Devine, Forscher, Austin, & Cox, 2012), which conceptualizes bias as a mental habit and lays out the steps needed to "break the bias habit." Specifically, once a person is motivated to act in less biased ways, breaking the bias habit involves 1) becoming aware of when one is vulnerable to unintentional bias, 2) understanding the consequences of unintentional bias, and 3) learning and practicing effective strategies to reduce the impact of unintentional bias.

Devine et al. (2012) operationalized the components of the habit-breaking model into the *prejudice habit-breaking intervention*, which is

☆ The datasets and supplemental analyses are publicly available online, at https://osf.io/9yt23/. The full analytic plan was preregistered at https://osf.io/jbs84/.

☆☆ Preparation of this article was supported by NIH grant 5R01GM111002-02, a WARF Professorship awarded to the first author, and an internal grant from the Office of the Provost of the University of Wisconsin-Madison awarded to the first and third authors.

* Corresponding author at: Department of Psychology, University of Wisconsin-Madison, 1202 W Johnson St., Madison, WI 53706, United States
E-mail address: pgdevine@wisc.edu (P.G. Devine).

http://dx.doi.org/10.1016/j.jesp.2017.07.002
Received 10 February 2017; Received in revised form 5 July 2017; Accepted 5 July 2017
Available online 17 August 2017

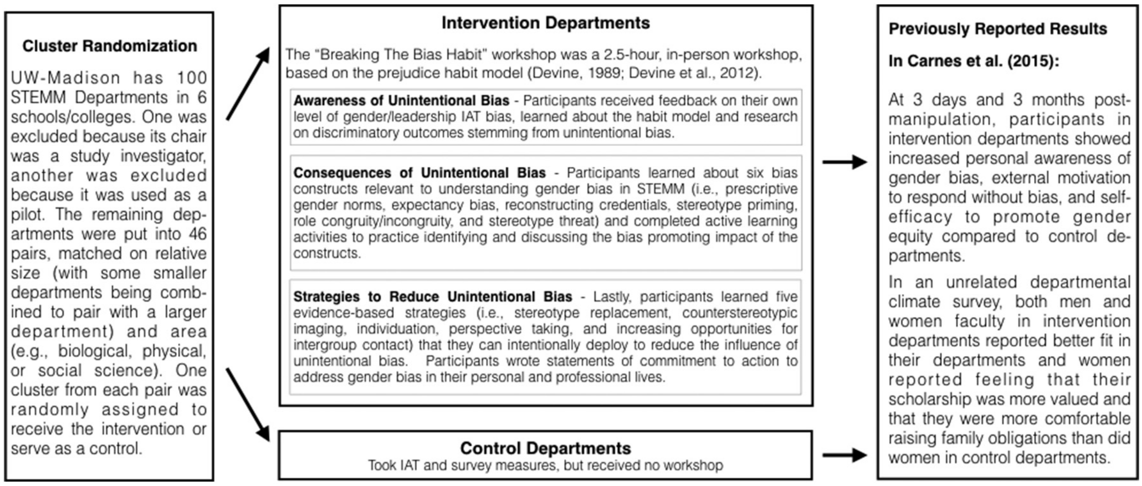

Fig. 1. The gender habit-breaking intervention. Study design, intervention components, and previously reported results.

thus far the only intervention experimentally shown to produce long-term changes in bias (Devine et al., 2012), with effects lasting at least 2 years post-manipulation (Forscher, Mitamura, Dix, Cox, & Devine, 2017). One iteration of this intervention approach is the *gender bias habit-breaking intervention* (Carnes et al., 2015), which focused specifically on gender bias in STEMM fields and was implemented in a 2.5 h workshop to individual departments.

The workshop (see Fig. 1 and Carnes et al., 2012) reviews the key components of the habit model (awareness, consequences, and strategies). To increase awareness, prior to the workshop participants completed and received feedback on a gender/leadership Implicit Association Test (IAT). The workshop opened with evidence of continuing gender bias in STEMM, including the underrepresentation of women in faculty and leadership positions and the potential adverse impact such biases for the overarching goals of advancing science, national health, and economic vitality. Attendees learned how unintentional bias function like habits, leading people to often respond in ways that contradict egalitarian values. They then learned about six "bias constructs" that represent common manifestations of gender bias generally and in STEMM more specifically (i.e., expectancy bias, prescriptive gender norms, role congruity/incongruity, stereotype priming, reconstructing credentials, and stereotype threat). To allow attendees to actively engage with the constructs and foster learning of new material, attendees next read and discussed case studies to practice identifying and examining the bias-promoting impact of the constructs. To promote efficacy to reduce bias, attendees learned five evidence-based strategies (i.e., stereotype replacement, counterstereotypic imaging, individuation, perspective taking, and increasing opportunities for intergroup contact) that have been shown to counteract unintentional bias (Devine et al., 2012); attendees were told that practicing the strategies would help them to break the gender bias habit. Attendees also wrote statements of commitment to action to address gender bias in their personal and professional lives, a strategy found to be effective in other contexts to promote behavioral change (Overton & MacVicar, 2008). By increasing attendees' understanding of unintentional gender bias and its adverse effects, we encouraged faculty to intentionally change their behavior to mitigate the impact of unintentional bias. We assumed that engaging faculty in this way would be the first step toward institutional transformation.

We tested the gender habit-breaking intervention's effectiveness in a large-scale cluster-randomized-controlled trial in 98 STEMM departments at the University of Wisconsin–Madison. Compared to control departments, intervention departments showed increases in personal awareness of gender bias and self-efficacy to promote gender equity three days and three months post-manipulation and increases in self-reported action to promote gender equity at the three-month assessment (Carnes et al., 2015). On an unrelated university climate survey, faculty in intervention departments reported feeling better fit in their departments, that their scholarship was more valued by their colleagues and that they were more comfortable raising family obligations than did faculty in control departments.

Although encouraging with regard to outcomes that would be expected to promote gender equity in STEMM, our previous results are exclusively self-report. To be impactful, the intervention must also produce changes in key behavioral outcomes related to reducing gender bias and STEMM. In the present work, we examine the impact of gender habit-breaking intervention on the gender of new faculty hires. We chose hiring patterns as our main outcome for a number of reasons. First, an effective intervention, ideally, would help reduce the underrepresentation of women in STEMM. Second, the intervention specifically discussed how bias can affect the likelihood of women being hired in STEMM (e.g., reconstructing credentials, role incongruity). Third, to the extent that unintentional gender bias contributes to the underrepresentation of women (see Moss-Racusin et al., 2012), participants' greater awareness of, and self-efficacy to overcome, unintentional bias as well as their written commitment to address gender bias should reduce the effects of unintentional bias on hiring, yielding more new women faculty hires. Fourth, hiring decisions are made by departments, not individuals, which is well-matched to the cluster-randomized design, in which departments were assigned to receive the intervention or serve as controls.[1] In prior tests of the impact of the habit-breaking intervention, outcomes were assessed at the individual level even when evaluated as the cluster level (Carnes et al., 2015; Devine et al., 2012; Forscher et al., 2017). In the present context, we explore the potential for the intervention to affect individuals in ways that may promote change in institutional level outcomes. Finally, to our knowledge, no past work has investigated the impact of a real-world intergroup bias intervention on this type of highly consequential outcome. We anticipated that, compared to control departments and intervention departments in the pre-manipulation period, only intervention departments in the post-intervention period would show greater gender balance in their new hires.

1. Method

The pre-registered analytic plan, dataset, and supplemental analyses are available at https://osf.io/9yt23/. All measures, manipulations, and exclusions are disclosed here and in Carnes et al. (2015). At the study

[1] Although all intervention department members were invited to attend, only a subset did.

outset, the 6 STEMM-focused schools/colleges at UW-Madison had 100 STEMM and our sample includes all but two of the departments. One department was excluded because its department chair was a study investigator, and the other was used as a pilot. Of the remaining 98, 6 small departments were combined into two clusters of three departments, yielding 92 clusters. These clusters were assigned to 46 pairs, matched on size, school/college, and disciplinary category. One member of each pair was randomly assigned to receive the intervention workshop and the other served as a control. Following completion of the two year workshop administration period, control departments were offered the workshop, but less than 2% of their faculty attended, enabling those departments to remain controls.

Our experimental approach has inherent strengths. First, our approach follows Moss-Racusin et al. (2014) recommendations for effective evidence-based interventions. Second, as a real-world randomized-controlled trial, it affords the opportunity to rigorously assess the causal impact of the gender habit-breaking intervention on hiring, an important outcome for addressing gender bias in STEMM. Third, our sample included all departments in the 6 STEMM-focused schools/colleges at UW-Madison. This strength, however, carries with it a specific potential limitation. Because random assignment occurs at the level of departmental clusters, the maximum sample size is 92 clusters, which places a hard upper limit on statistical power to detect effects.

We compared hiring and attrition rates during the two years before the workshops began and the two years after the workshops were completed, thereby keeping all departments equal with regard to university-level factors (e.g., budgetary concerns) that could affect hiring. Using annual human resources records, faculty members who were not in the previous year's database were counted as new hires. Faculty members who were in the previous year's but not the current year's database were counted as leaving the department.

2. Results

2.1. Descriptive hiring rates

In the pre-manipulation period, control departments hired 109 faculty (33% women) and intervention departments hired 85 faculty (32% women). During the post-manipulation period control departments made 113 hires (32% women) and intervention departments made 101 hires (47% women).[2] To protect against potential spurious effects arising from collapsing across separate, independent hiring units (i.e., Simpson's Paradox; Pearl, 2000; Simpson, 1951), the formal test of our hypothesis treats departmental clusters as the unit of analysis.

2.2. Preregistered analyses

Analyses were conducted using Generalized Linear Mixed Effects Models (GLMEMs) with a logit link in the binomial family with the bobyqa optimizer in the lme4 package in R (Bates et al., 2015). Each analysis included random intercepts for each departmental cluster and each matched pair of clusters. The number of female hires was weighted by the total number of hires within the cluster, so the outcomes could be interpreted as proportions. In each model, we tested the interaction between an indicator for condition (control = -0.5; intervention = 0.5) and an indicator for time period (0 = pre; 1 = post). We used each coefficient to calculate an odds ratio (*OR*) and tested whether its profile likelihood 95% confidence interval overlapped with 1. For *OR*s, a value of ~1.5 is considered small, ~3.5 is considered medium, and ~9 is considered large (Wuensch, 2009). Haddock and colleagues provide an equation to convert *OR*s to an equivalent of Cohen's *d*, which we report alongside the *OR*s (Haddock, Rinkdskopf, & Shadish, 1998).

As shown in Table 1,[3] there was modest evidence that, whereas the proportion of women hired by control departments remained stable over time, the proportion of women hired by intervention departments increased, $OR = 2.23$, $\chi^2(1, n = 81) = 3.25$, $p = 0.07$, 95% CI = [0.94, 5.41], $d_{OR} = 0.34$. Descriptively, in the pre-manipulation period, control and intervention departments did not differ in the proportion of new female faculty hires, $OR = 0.95$, $p = 0.87$, 95% CI = [0.50, 1.80], $d_{OR} = -0.02$, and intervention departments hired a higher proportion of women in the post-manipulation period than control departments $OR = 2.12$, $p = 0.02$, 95% CI = [1.16, 3.95], $d_{OR} = 0.32$. In addition, whereas the proportion of women hired by control departments remained stable over time $OR = 0.82$, $p = 0.53$, 95% CI = [0.45, 1.50], $d_{OR} = -0.08$, the proportion of women hired by intervention departments increased from the pre- to post-manipulation period, $OR = 1.84$, $p = 0.06$, 95% CI = [0.98, 3.50], $d_{OR} = 0.26$.

There was no evidence of a change in the overall gender composition of the faculty post-manipulation. The proportion of female faculty who left the departments increased in the intervention departments compared to control departments, $OR = 3.03$, $\chi^2(1, n = 87) = 4.04$, $p = 0.04$, 95% CI = [1.04, 9.11], $d_{OR} = 0.47$, which we explore next in more detail.

2.3. Exploratory "revolving door" analyses

Taken together, the hiring and attrition patterns raise the possibility that any apparent increase in new women hires may reflect a "revolving door" whereby female faculty leave and departments merely replace them. We tested the "revolving door" account by separately estimating the change in the numbers, rather than the proportion, of hires and attrition for men and women using GLMEMs with log links from the Poisson family. As shown in Table 1, the increase in intervention departments' proportions of female attrition was driven by decreases in the number of men who left intervention departments relative to control departments, $RR = 0.64$, $\chi^2(1, n = 89) = 3.62$, $p = 0.06$, 95% CI = [0.40, 1.01], not increases in the number of women who left, $RR = 1.68$, $\chi^2(1, n = 89) = 1.28$, $p = 0.26$, 95% CI = [0.68, 4.18]. Given previously-reported patterns showing the gender habit-breaking intervention improved climate for both women *and* men (Carnes et al., 2015), it is reasonable that men in intervention departments may have been less likely to leave.

3. Discussion

Our findings are promising with regard to improving the representation of women in STEMM disciplines. According to our preregistered GLMEM's estimate, intervention departments hired 18 percentage points more women in the post- than pre-intervention period. Control departments did not vary in their hiring of women over time. Pre-intervention, hires in control and intervention departments substantially favored men, but after the manipulation, new hires in intervention departments were gender balanced. This gender-balanced hiring is what one would expect if there are equal numbers of qualified men and women applicants. Though increased hiring of women did not achieve conventional levels of statistical significance, our study has a hard upper limit on statistical power. We hasten to add, however, that statistical certainty is only one criterion against which to judge the

[2] The overall number of hires by cluster in the pre-intervention period was marginally different, $\chi^2(1, n = 92) = 2.95$, $p = 0.09$, but intervention and control departments hired approximately the same number of new faculty in the post-intervention period, $\chi^2(1, n = 92) = 0.67$, $p = 0.41$. Although the control departments hired somewhat more faculty overall in the pre-manipulation period, the greater number of hires did not yield a greater gender balance in hiring. Because our analyses either operate on proportions or are weighted by the total number of hires, we do not think this pre-manipulation difference affects the interpretation of the hiring proportions.

[3] In addition to gender, our preregistered analyses examined effects related to underrepresented minorities (URMs), which are also reported in Table 1. The overall number of URMs was small, and there was no evidence of changes in hiring or attrition of URMs.

Table 1
The GLMEM-estimated proportions of women and underrepresented minorities and the GLMEM-estimated numbers of women and men among new hires, current faculty, and faculty who left the department broken down by time period and condition. The "Δ" columns represent the pairwise comparison in odds rations (ORs; for proportions) or relative ratios (RRs; for numbers).

		Estimates				Time × condition			Control vs. intervention						Pre vs. post					
		Control		Intervention		Interaction			Pre			Post			Control			Intervention		
		Pre	Post	Pre	Post	Int.	95% CI	p	Δ	95% CI	p	Δ	95% CI	p	Δ	95% CI	p	Δ	95% CI	p
Proportion of women	Hires	0.34	0.29	0.32	0.47	2.23	[0.94, 5.41]	0.07	0.95	[0.50, 1.80]	0.87	2.12	[1.16, 3.95]	0.02	0.82	[0.45, 1.50]	0.53	1.84	[0.98, 3.50]	0.06
	Faculty	0.21	0.23	0.22	0.24	1.02	[0.71, 1.47]	0.89	1.02	[0.75, 1.39]	0.89	1.05	[0.77, 1.42]	0.77	1.12	[0.87, 1.44]	0.38	1.14	[0.88, 1.49]	0.31
	Attrition	0.20	0.19	0.12	0.27	3.03	[1.04, 9.11]	0.04	0.53	[0.22, 1.19]	0.13	1.59	[0.67, 3.82]	0.29	0.93	[0.46, 1.86]	0.83	2.81	[1.24, 6.60]	0.01
Proportion of URMs	Hires	0.08	0.06	0.07	0.10	1.96	[0.47, 8.59]	0.36	0.93	[0.31, 2.61]	0.88	1.81	[0.67, 5.14]	0.25	0.74	[0.26, 2.01]	0.56	1.45	[0.53, 4.27]	0.48
	Faculty	0.03	0.04	0.03	0.04	1.21	[0.61, 2.41]	0.97	0.93	[0.46, 1.84]	0.83	1.12	[0.56, 2.20]	0.85	1.01	[0.62, 1.64]	0.97	1.22	[0.75, 1.99]	0.43
	Attrition	0.03	0.05	0.02	0.04	1.62	[0.26, 11.28]	0.61	0.53	[0.10, 2.30]	0.41	0.86	[0.18, 3.69]	0.41	1.48	[0.48, 4.78]	0.49	2.41	[0.56, 12.10]	0.25
Numbers of women	Hires	0.69	0.65	0.53	0.91	1.80	[0.93, 3.54]	0.08	0.77	[0.46, 1.28]	0.31	1.39	[0.88, 2.21]	0.16	0.94	[0.59, 1.51]	0.81	1.70	[1.07, 2.77]	0.03
	Faculty	2.93	3.10	2.77	3.02	1.03	[0.76, 1.39]	0.85	0.95	[0.73, 1.24]	0.68	0.98	[0.75, 1.27]	0.85	1.06	[0.86, 1.30]	0.59	1.09	[0.88, 1.36]	0.43
	Attrition	0.43	0.35	0.24	0.32	1.68	[0.68, 4.18]	0.26	0.56	[0.26, 1.16]	0.12	0.93	[0.45, 1.92]	0.85	0.81	[0.45, 1.43]	0.47	1.36	[0.68, 2.76]	0.39
Numbers of men	Hires	1.27	1.34	1.00	0.94	0.88	[0.54, 1.44]	0.62	0.79	[0.52, 1.18]	0.25	0.70	[0.46, 1.05]	0.08	1.05	[0.52, 1.18]	0.74	0.93	[0.54, 1.35]	0.71
	Faculty	10.13	9.83	9.51	9.28	1.00	[0.85, 1.19]	0.96	0.94	[0.77, 1.15]	0.54	0.94	[0.77, 1.16]	0.57	0.97	[0.87, 1.09]	0.62	0.98	[0.86, 1.10]	0.67
	Attrition	1.45	1.36	1.36	0.82	0.64	[0.40, 1.01]	0.06	0.94	[0.67, 1.32]	0.72	0.60	[0.41, 0.89]	0.01	0.94	[0.70, 1.27]	0.69	0.60	[0.42, 0.86]	< 0.01

importance of findings (Ross & Nisbett, 2011). Ross and Nisbett also highlight the importance of pragmatic criteria, and we contend that the shift observed in hiring of women has undeniable practical significance for the long-term goal of achieving gender equity in STEMM.

We can only speculate about the processes that may have produced gender-balanced hiring. It is possible, for example, that faculty who attended the workshop became more concerned about gender discrimination (Forscher et al., 2017), which may have led them to be more active in hiring committees and more proactive in considering and advocating for female applicants (Bardi & Schwartz, 2003; Krosnick, 1988). The faculty may also have implemented strategies to circumvent their own gender biases (Devine et al., 2012) or identified and labeled common manifestations of gender bias in others, setting the stage for constructive conversations with colleagues about gender equity (Ashburn-Nardo, Morris, & Goodwin, 2008; Forscher et al., 2017; Mitamura, Erickson, & Devine, 2017; Nonaka, 1994). These processes may have caused intervention departments to seek out and make more offers to women. Alternatively, given that intervention departments appear to have better climates for women (and men), perhaps they were more effective at successfully recruiting women once an offer was extended. To the extent that these or other processes demonstrate an institutional commitment to the professional success of female faculty, they could have recursive and synergistic effects on future hiring and retention, with the potential to effect broader institutional change. Investigating these possibilities is a high priority for future work.

Though we are cautiously optimistic about our findings, we acknowledge that the marginal statistical significance of the hiring effect does not permit a high degree of certainty in our intervention's influence on hiring. Moreover, hiring is but the first step on the long journey to achieving gender equity in STEMM. As yet, there is no evidence that the intervention caused a change in the overall gender composition of experimental departments. Such a change would likely require hiring changes that endure much longer than two years. After hiring has occurred, gender equity can only truly be achieved if women thrive in their departments, achieve tenure, and ascend to leadership positions.

Our confidence in these findings, however, is bolstered by the strength of the cluster-randomized controlled experimental design and the long-term, longitudinal assessment of the intervention's impact (Moss-Racusin et al., 2014). Though conducting comprehensive, longitudinal, theoretically-derived intervention work is an enormous undertaking, our study reveals the potential payoffs of such large-scale efforts. Moreover, we believe this type of approach is necessary to stem the tide of ineffective and sometimes harmful bias-reducing approaches based on intuition or wishful thinking. Translating psychological research into application in the form of evidence-based interventions is essential to fulfill the promise of psychological science as a force to improve people's lives and society.

Author contributions

P. G. Devine and P. S. Forscher conceived the study concept. J. Sheridan retrieved the data from the University of Wisconsin-Madison Human Resources records. P. S. Forscher and W. T. L. Cox analyzed the data under the supervision of P. G. Devine. W. T. L. Cox, P. G. Devine, and P. S. Forscher planned and drafted the paper. M. Carnes was PI of NIH R01 GM088477 that funded the study on which the current research is based. All authors provided edits and approved the final version of the manuscript for submission.

Open practices

The experiment in this article earned Open Materials, Open Data, and Preregistered badges for transparent practices. Materials, data, and preregistration information for the experiment are available at https://osf.io/9yt23/.

References

Apfelbaum, E. P., Norton, M. I., & Sommers, S. R. (2012). Racial colorblindness: Emergence, practice, and implications. *Current Directions in Psychological Science*, *21*, 205–209.

Ashburn-Nardo, L., Morris, K. A., & Goodwin, S. A. (2008). The confronting prejudiced responses (CPR) model: Applying CPR in organizations. *Academy of Management Learning & Education*, *7*, 332–342.

Bardi, A., & Schwartz, S. H. (2003). Values and behavior: Strength and structure of relations. *Personality and Social Psychology Bulletin*, *29*, 1207–1220.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., & Singmann, H. (2015). lme4: Linear mixed-effects models using Eigen and S4, 2014. *R package version*, *1*(4).

Carnes, M., Devine, P. G., Isaac, C., Manwell, L. B., Ford, C. E., Byars-Winston, A., ... Sheridan, J. (2012). Promoting institutional change through bias literacy. *Journal of Diversity in Higher Education*, *5*, 63.

Carnes, M., Devine, P. G., Manwell, L. B., Byars-Winston, A., Fine, E., Ford, C. E., ... Sherridan, J. (2015). Effect of an intervention to break the gender bias habit: A cluster randomized, controlled trial. *Academic Medicine*, *90*, 221–230.

Corrice, A. (2009). Unconscious bias in faculty and leadership recruitment: A literature review. *AAMC Analysis in Brief*, *9*, 1–2.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit racial bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology*, *48*, 1268–1278.

Dobbin, F., & Kalev, A. (2013). The origins and effects of corporate diversity programs. In Q. Q. M. Robesson (Ed.), *Oxford Handbook of Diversity and Work* (pp. 253–281). New York: Oxford Press.

Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T. L., & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology*, *72*, 133–146.

Haddock, C. K., Rinkdskopf, D., & Shadish, C. (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. *Psychological Methods*, *3*, 339–353.

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. 1111 Sixteenth Street NW, Washington, DC 20036: American Association of University Women.

Krosnick, J. A. (1988). The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluations, and voting behavior. *Journal of Personality and Social Psychology*, *55*, 196.

Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages how motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, *22*, 1472–1477.

Mitamura, C., Erickson, L., & Devine, P. G. (2017). Value-based standards guide sexism inferences for self and others. *Journal of Experimental Social Psychology*, *72*, 101–117.

Mitchneck, B., Smith, J. L., & Latimer, M. (2016). A recipe for change: Creating a more inclusive academy. *Science*, *352*, 148–149.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*. *109*. *Proceedings of the National Academy of Sciences* (pp. 16474–16479).

Moss-Racusin, C. A., van der Toorn, J., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2014). Scientific diversity interventions. *Science*, *343*, 615–616.

National Academy of Sciences, National Academy of Engineering, Institute of Medicine of the National Academies (2006). *Beyond biases and barriers: Fulfilling the potential of women in academic science and engineering*. Washington, DC: National Academies Press.

National Science Foundation (2007). *Women, minorities, and persons with disabilities in science and engineering*. (NSF 07–315).

National Science Foundation (2014). *Pathways to broadening participation in response to the CEOSE 2011–2012 recommendation*.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, *5*, 14–37.

Overton, G. K., & MacVicar, R. (2008). Requesting a commitment to change: Conditions that produce behavioral or attitudinal commitment. *Journal of Continuing Education in the Health Professions*, *28*, 60–66.

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology*, *60*, 339–367.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Ross, L., & Nisbett, R. E. (2011). *The person and the situation: Perspectives of social psychology*. London, UK: Pinter & Martin Publishers.

Sevo, R., & Chubin, D. E. (2008). *Bias literacy: A review of concepts in research on discrimination*. American Association for the Advancement of Science Center for Science & Engineering Capacity.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B*, *13*, 238–241.

Valantine, H. A., & Collins, F. S. (2015). National Institutes of Health addresses the science of diversity. *Proceedings of the National Academy of Sciences*. *112*. *Proceedings of the National Academy of Sciences* (pp. 12240–12242).

Wuensch, K. (2009). Rules of thumb on magnitudes of effect sizes. Retrieved from http://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize.