Increasing the perceived malleability of gender bias using a modified Video Intervention for Diversity in STEM (VIDS)

Article *in* Group Processes & Intergroup Relations · August 2018

DOI: 10.1177/1368430218755923

CITATIONS
2

READS
106

8 authors, including:

Corinne A. Moss-Racusin
Skidmore College
48 PUBLICATIONS **2,227** CITATIONS
SEE PROFILE

April H. Bailey
Yale University
14 PUBLICATIONS **31** CITATIONS
SEE PROFILE

Jo Handelsman
University of Wisconsin–Madison
410 PUBLICATIONS **22,664** CITATIONS
SEE PROFILE

Group Processes & Intergroup Relations

Article

Increasing the perceived malleability of gender bias using a modified Video Intervention for Diversity in STEM (VIDS)

Group Processes & Intergroup Relations
2018, Vol. 21(5) 788–809

DOI: 10.1177/1368430218755923
journals.sagepub.com/home/gpi
SAGE

Erin P. Hennes,[1*] Evava S. Pietri,[2*] Corinne A. Moss-Racusin,[3] Katherine A. Mason,[1] John F. Dovidio,[4] Victoria L. Brescoll,[4] April H. Bailey,[4] and Jo Handelsman[5]

Abstract

Scholars are increasingly responding to calls for interventions to address persistent gender disparities in the sciences. Yet, interventions that emphasize the pervasiveness of bias may inadvertently damage efficacy to confront sexism by creating the perception that bias is immutable. We examined this possibility in the context of a successful bias literacy program, Video Interventions for Diversity in STEM (VIDS; Moss-Racusin et al., in press). In two studies with working adults from the general public ($N = 343$) and science faculty ($N = 149$), we modified VIDS by developing a module (UNITE) that offers tools for addressing bias and promotes the mindset that bias is malleable. VIDS alone was sufficient to increase awareness of bias, reduce sexism, and improve bias identification. However, UNITE buffered against perceptions that bias is immutable and restored self-efficacy to address bias. We conclude that interventions must aim not only to increase bias literacy but also offer concrete tools and avoid implying that these problems are insurmountable.

Keywords

diversity intervention, gender bias, mindset, self-efficacy

Paper received 1 April 2017; revised version accepted 5 January 2018.

[1]Purdue University, USA
[2]Indiana University–Purdue University Indianapolis, USA
[3]Skidmore College, USA
[4]Yale University, USA
[5]University of Wisconsin–Madison, USA

Corresponding authors:
Erin P. Hennes, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907, USA.
Email: ehennes@purdue.edu

Evava S. Pietri, IUPUI, 402 North Blackford, Indianapolis, IN 46202, USA.
Email: epietri@iupui.edu

In recent years, the underrepresentation of women in many areas of science, technology, engineering, and math (STEM) has received increased attention (Ceci & Williams, 2011; National Science Foundation [NSF], 2017). Several studies have suggested that one cause of this disparity is gender discrimination (e.g., Milkman, Akinola, & Chugh, 2015; Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012), leading practitioners to respond to calls for interventions to improve gender equity (e.g., Penner, 2015; "Sexism Has No Place," 2015). Unfortunately, many of these diversity interventions have proven unsuccessful, and have even led to ironic prejudice or backlash effects (e.g., Bezrukova, Spell, Perry, & Jehn, 2016; Dobbin & Kalev, 2016; Duguid & Thomas-Hunt, 2015; Legault, Gutsell, & Inzlicht, 2011), or they have not been subjected to evaluation through randomized controlled trials (see Moss-Racusin et al., 2014; Paluck & Green, 2009, for reviews). Even those that effectively lead to bias reduction may be difficult to implement on a large scale (e.g., because they require trained facilitators; Carnes et al., 2015; Zawadzki, Danube, & Shields, 2012) or may have inadvertent costs (e.g., Liben & Coyle, 2014). Moreover, some scholars have cautioned that efforts to improve parity may be misguided (e.g., Ceci, Ginther, Kahn, & Williams, 2014) or may lead to oversensitivity to bias when none actually exists (Haidt, 2017). Thus, it is critical that diversity interventions effectively reduce bias, are suitable for wide-scale implementation, and minimize unintended consequences (Moss-Racusin et al., 2014).

One potential unintended consequence of diversity interventions is a resulting belief that gender bias is insurmountable and immutable (i.e., a fixed rather than a growth mindset; Bandura, 2004; Carnes et al., 2012; Carr, Dweck, & Pauker, 2012). Indeed, health and organizational behavior research suggests that exposing participants to evidence of a serious problem, such as the existence of gender bias in STEM, without providing concrete strategies for combating the problem, may inadvertently lead participants to feel overwhelmed rather than efficacious and to believe that such problems are normative and immutable (Bandura, 2004; Block & Keller, 1995; Duguid & Thomas-Hunt, 2015; Good & Abraham, 2011; Stephenson & Witte, 1998). We explored this possibility in the context of a previously validated intervention, Video Interventions for Diversity in STEM (VIDS), which exposes participants to empirical psychological evidence of gender bias in a compelling media-based format. VIDS has been found, in randomized controlled trials, to reduce gender bias and increase bias awareness but without leading to bias hypersensitivity (Moss-Racusin et al., in press; Pietri et al., 2017).

The aim of the current research was to examine and counteract the potential impact of bias awareness training on self-efficacy and growth mindset about bias reduction. To do so, we developed a novel additional component to VIDS (referred to as "UNITE" in reference to an acronym used in the intervention), which provides specific strategies for combating bias and draws from mindset theory to promote the idea that bias can be changed (Dweck, 2000). Our goal was to develop an enhanced intervention to not only reduce individuals' own bias but to empower them to confront prejudice in others and promote equity in STEM.

The Importance of Reducing Gender Bias in STEM

The underrepresentation of women in STEM does not mean that women are less capable than are men in these domains (e.g., Lindberg, Hyde, Petersen, & Linn, 2010). Rather, this disparity may, in part, result from stereotypes associating scientists with traditionally male traits (e.g., assertive, aggressive; Diekman, Brown, Johnston, & Clark, 2010; Nosek, Banaji, & Greenwald, 2002; Nosek et al., 2007). Such stereotypes can lead to the unequal treatment of men and women in STEM fields (Bilimoria & Liang, 2013; Cejka & Eagly, 1999; Milkman et al., 2015; Moss-Racusin et al., 2012; Renzulli, Grant, & Kathuria, 2006; Wright et al., 2003; but see Williams & Ceci, 2015) and can create a hostile and unwelcoming

environment for women (Cheryan, 2012; Cheryan, Plaut, Davies, & Steele, 2009). Indeed, both men and women faculty members show biases favoring men (Milkman et al., 2015; Moss-Racusin et al., 2012), even those who explicitly value gender diversity (Nosek et al., 2007). Further, the subtlety of gender bias may prevent individuals from recognizing that women experience discrimination (Glick & Fiske, 2001; Rudman & Glick, 2008; Swim, Hyers, Cohen, & Ferguson, 2001). Nevertheless, these subtle biases negatively impact women's career trajectories and wellbeing (Cortina, 2008; Settles, Cortina, Buchanan, & Miner, 2013). Consequently, failure to notice and address sexism may result in its continued harmful impact on women's progress in STEM fields.

Bias Interventions

Because gender bias detrimentally affects women in STEM, it is necessary to develop effective interventions that reduce sexism to guarantee equal opportunity in these fields. Although effective theory-based gender bias interventions are rare, they do exist. One example taught participants about gender inequity through experiential learning using an interactive game that indirectly exposed participants to the challenges women encounter in the workplace (Zawadzki et al., 2012). In another, presenters described the causes and consequences of gender bias in STEM and provided strategies for reducing biases and promoting fair treatment in participants' departments (Carnes et al., 2015). In both cases, the interventions led to reduced gender bias.

Despite these promising results, workshop interventions are time consuming for participants and require trained facilitators, making them difficult to execute on a broad scale. To address this limitation, a new intervention was recently developed—Video Interventions for Diversity in STEM (VIDS). VIDS consists of two sets of professionally produced, scripted videos that present the findings of social psychological research on gender bias. These videos utilize a combination of *narrative videos*, which follow the storylines of characters who have been negatively affected by gender bias in the sciences, and *expert interview videos*, which portray a psychology professor describing research on gender bias (see Moss-Racusin et al., in press; Pietri et al., 2017, for a detailed discussion of VIDS). Supporting the promise of media-based interventions for positive attitude change (e.g., Paluck, 2009; Paluck & Green, 2009), in randomized controlled trials the intervention demonstrated substantial improvements on awareness of gender bias in STEM and modern sexism among both science faculty and members of the general population. Importantly, VIDS increased participants' ability to identify unfair treatment across a variety of hypothetical situations but did not lead participants to claim bias indiscriminately (Pietri et al., 2017; cf. Haidt, 2017). Reductions in gender bias persisted for at least one week (Moss-Racusin et al., in press). Thus, VIDS appears to be an effective and scalable intervention for reducing gender bias in STEM.

Promoting Self-Efficacy to Address Gender Bias

Although VIDS effectively addresses bias literacy, improves bias identification, and reduces gender bias itself, there is little evidence that it leads to efficacy to address bias. Indeed, the videos offered no concrete strategies for combating bias (Carnes et al., 2015). Promoting self-efficacy is critical because people may not address harmful sexist actions if they do not feel that they have the ability or tools to change their, or others', behavior (Bandura, 1977; Carnes et al., 2015). For example, research on health messages has found that merely providing information about a given problem (e.g., risk factors for skin cancer) may result in message recipients feeling hopeless and discounting the message (Bandura, 2004; Floyd, Prentice-Dunn, & Rogers, 2000). However, providing information about specific actions to prevent harmful outcomes (e.g., wearing sunscreen) encourages self-efficacy (e.g., beliefs about personally preventing skin cancer) and healthy behaviors (e.g., regular cancer screenings; Block & Keller, 1995; Good & Abraham, 2011; Stephenson &

Witte, 1998). Using the health messages literature as a model, some diversity researchers have emphasized the importance of not only improving attitudes, but also stimulating feelings of self-efficacy to combat bias (e.g., Zawadzki et al., 2012) by providing strategies for doing so (Carnes et al., 2012; Carnes et al., 2015).

Promoting a Growth Mindset About Gender Bias

In addition to harming self-efficacy, VIDS has the potential to stimulate a fixed mindset about bias (i.e., perceptions that bias is stable and cannot be reduced). Indeed, when individuals learn about pervasive and persistent biases, they tend to believe that biases are fixed and unchangeable (Carr et al., 2012). Awareness of the pervasiveness of bias may also lead to the perception that prejudice is socially normative, further inhibiting both efficacy and motivation to reduce gender bias in oneself or others (Duguid & Thomas-Hunt, 2015). Research indicates that a growth mindset (i.e., the perception that bias can be changed) may be critical for addressing bias. For example, people who have a growth mindset about bias reduction are more likely to participate in strategies that reduce their personal biases (e.g., participating in a bias-reduction training or taking the perspective of a stigmatized group; Carr et al., 2012; Neel & Shapiro, 2012), and are also more likely to combat unfair treatment by others (e.g., confronting an individual who makes discriminatory comments; Rattan & Dweck, 2010). Although VIDS addressed shortcomings of previous workshop-based interventions in that it is easy to administer on a broad scale, it is imperative to further develop this program by encouraging self-efficacy and promoting the message that sexism can be overcome.

Current Research

To encourage self-efficacy and a growth mindset about bias reduction, we created an educational module called UNITE to supplement the VIDS intervention. UNITE begins by providing empirical information about gender bias in the workplace. It then provides detailed scientific and anecdotal evidence, based on mindset theory, that gender bias is not fixed, culminating in a step-by-step guide (using an acronym spelling the word "UNITE") for promoting gender equity using empirically based strategies.

We anticipated that VIDS would continue to function as an efficacious intervention to promote bias literacy and reduce sexism. Importantly, consistent with Pietri et al. (2017), we expected VIDS to improve individuals' ability to recognize bias, but *not* lead participants to claim bias in situations in which treatment was actually fair (Hypothesis 1). In contrast to these positive effects, we also expected VIDS to fail to improve—or possibly even harm—self-efficacy to address gender bias and perceptions that bias can be changed (Hypothesis 2). We predicted that the UNITE module would address these shortcomings by reinforcing a growth mindset and feelings of self-efficacy to combat gender bias in STEM (Hypothesis 3). Consequently, we aimed to show that combining VIDS with UNITE creates a powerful and easy to implement intervention, and demonstrate the general benefits associated with developing multiple component interventions that target multiple outcomes. We explored this possibility with working adults from the general population (Experiment 1) and longitudinally with STEM faculty (Experiment 2).

Experiment 1

Experiment 1 aimed, first, to replicate previous findings that VIDS successfully reduces gender bias (Moss-Racusin et al., in press), improves awareness of bias, and improves the ability to identify bias when it occurs without leading to indiscriminate claims of sexism (Pietri et al., 2017). Novel to the current research, we also sought to examine whether VIDS was sufficient to increase self-efficacy and growth mindset, or whether supplementing VIDS with our newly developed UNITE module would lead to added benefits, perhaps even buffering against unintended negative consequences of VIDS. To

examine these questions, we randomly assigned members of the general public to condition in a 2 (Video intervention: VIDS vs. control) x 2 (module intervention: UNITE vs. control) between-participants design.

Method and Procedure

Participants. Participants completed the experiment, advertised as a study of "Impressions of short movies and modules," in exchange for $5.00 on Amazon's Mechanical Turk (MTurk). Because the study concerned bias in the workplace, we restricted analyses to participants who were employed full-time outside the home.[1] Based on previous observation that slightly more than half of MTurk workers fulfill this criterion, we recruited 610 participants with the aim of a final sample of approximately 350 employed participants. This estimate was consistent with an a priori power analysis based on the effect size of $d = 0.31$ of the impact of VIDS on gender bias reduction (the smaller of the previously observed effects of VIDS) reported by Moss-Racusin et al. (in press). Our final sample included 343 participants (37% women; 75% White). Participants ranged in age from 20 to 70 ($M = 35.24$, $SD = 9.93$) and the plurality (63%) had a college degree. On a scale of 1 (*strongly liberal*) to 9 (*strongly conservative*), participants' average political orientation was 4.27 ($SD = 2.35$).

Materials

VIDS. Participants were exposed to four videos selected from the library of videos developed and described in detail by Moss-Racusin et al. (in press). Participants in the intervention (VIDS) condition were exposed to two *narrative* videos and two corresponding *expert interview* videos. Participants varied randomly in whether they viewed the narrative video or the corresponding expert interview video first (order did not impact upon any results). The narratives were stories illustrating the empirical results of selected published papers on gender bias and stereotypes, and were written by a professional playwright to ensure that the script was entertaining, emotionally engaging, and transporting. The narrative videos featured professional actors playing science professors, graduate students, and laboratory technicians. The expert interviews described the same psychological research displayed in the narrative films, but in a straightforward, fact-based manner during an interview with a psychology professor (the expert; all videos are available at https://academics.skidmore.edu/blogs/vids/). Participants in the control video condition were exposed to four 4- to 6-minute clips from existing science documentaries, also chosen from the library of videos used in previous VIDS trials (Pietri et al., 2017).

UNITE. Following the videos, participants completed one of two modules. Both modules were created using Microsoft PowerPoint and consisted of information provided via text, graphs, and images. The module advanced automatically, ensuring that exposure time to the module was fixed. Because not all participants were exposed to VIDS, the UNITE module began with a brief evidence-based review of gender disparities in the workforce with attendant citations. The module then communicated that bias is malleable, and that if people are motivated, they have the ability to decrease their biases and the biases of those around them (Rattan & Dweck, 2010). This information was again accompanied by graphical representations of evidence and citations. Additionally, UNITE provided examples of individuals who have improved their implicit biases and promoted equity in their fields. UNITE then provided concrete evidence-based recommendations for creating an inclusive workplace. It did so by giving "tips" about how to decrease bias and promote equity, which made up the acronym "UNITE" (i.e., Underscore effective diversity training, Notice and correct for your implicit biases, Include inclusive pictures and language, Take time to mentor your fellow employees, Emphasize that employees can and will improve). Each tip cited research findings illustrating the behavior's effectiveness for reducing bias and included "dos" and "don'ts" (e.g. "Don't: Suggest that people who succeed have a special talent that

Table 1. Descriptive and reliability statistics (Experiments 1 and 2).

Variable	Experiment 1			Experiment 2										
				Baseline			Postintervention			Follow-up			R_{1F}	R_C
	M	*SD*	α	*M*	*SD*	α	*M*	*SD*	α	*M*	*SD*	α		
Awareness of gender bias in STEM	3.69	0.83	.89	3.64	0.63	.84	3.81	0.64	.87	3.75	0.65	.86	.88	.58
Gender bias	2.35	0.87	.90	1.99	0.55	.81	1.95	0.54	.83	1.89	0.54	.82	.87	.41
Bias identification	3.68	0.71	.78	-	-	-	-	-	-	-	-	-	-	-
Bias false identification	1.72	1.03	.83	-	-	-	-	-	-	-	-	-	-	-
Growth mindset	4.02	0.91	.90	4.10	0.63	.84	4.13	0.71	.86	4.19	0.78	.92	.83	.79
Self-efficacy	4.08	0.63	.79	3.66	0.62	.55	3.87	0.64	.66	3.83	0.71	.73	.76	.38

Note. *M* = grand mean; *SD* = standard deviation, α = Cronbach's alpha estimate of reliability, R_{1F} = generalized reliability; R_C = reliability of change (Cranford et al., 2006).

they are 'born' with"). The module aimed, first, to improve beliefs that gender bias can be changed, and second, to offer concrete tools in order to foster self-efficacy to confront bias in themselves and others.

The control module discussed the benefits of improving engagement in the workplace. It encouraged a growth mindset regarding employee engagement (e.g., communicating that when people are motivated, they have the power to improve their engagement and the engagement of those around them) and provided a set of "tips," as well as "dos" and "don'ts" about how to improve engagement in the workplace (all modules from Experiments 1 and 2 are available at https://academics.skidmore.edu/blogs/vids/unite-module/).

Measures. Descriptive and reliability statistics for all scales (in both experiments) are presented in Table 1.

Awareness of gender bias in the workplace. After watching one of the two sets of videos (VIDS vs. control) and one of two modules (UNITE vs. control), participants completed a series of measures. To assess awareness of gender bias in the workplace, we adapted a scale from previous research (Pietri et al., 2017). Participants responded to nine items on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*) including "In my opinion, working women often are not taken as seriously as their men colleagues." Items were identical to those of Pietri et al. (2017), with the exception that wording was altered to refer to the general, rather than STEM-specific, workplace. Items were averaged to create the awareness of bias index (α = .89).

Gender bias. As in prior research assessing VIDS (Moss-Racusin et al., in press), gender bias was assessed using the Modern Sexism Scale, which is a well-validated instrument frequently employed to measure subtle, contemporary forms of bias against women (Swim, Aikin, Hall, & Hunter, 1995). Participants responded to eight items on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*), including "Discrimination against women is no longer a problem in the United States." Items were averaged to create the modern sexism scale (α = .90).

Bias identification and false identification. To examine whether VIDS improved participants' ability to recognize bias, and to ensure that the intervention did not lead them to claim bias indiscriminately, we exposed participants to eight brief vignettes describing encounters between two people working in STEM fields (Pietri et al., 2017). Five of the situations indicated gender bias (e.g., a coworker is irritated with a research assistant

Table 2. Bivariate correlation matrix (Experiment 1).

Variable	Awareness of bias	Gender bias	Bias identification	Bias false identification	Growth mindset	Self-efficacy	Conservatism	Age
Awareness of bias								
Gender bias	−.79***							
Bias identification	.62***	−.62***						
Bias false identification	−.12*	.21***	−.07					
Growth mindset	.35***	−.42***	.30***	−.34***				
Self-efficacy	.23***	−.22***	.29***	−.12*	.31***			
Conservatism	−.34***	.46***	−.34***	.04	−.23***	−.08		
Age	.11*	−.10†	.08	−.12*	.08	.02	.18**	
Gender (W = 0, M = 1)	−.30***	.28***	−.22***	.05	−.06	−.03	.18**	−.21***

†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

for becoming pregnant, and when reminded that another assistant became a father the previous year, the coworker says that, "everyone knows it's different with mothers"), whereas three did not (e.g., a coworker is thinking about firing one of his employees because she is consistently late to work, has left on many occasions without asking permission, and missed work without calling in first). After reading a brief vignette, participants rated their level of agreement (1 = *strongly disagree*, 5 = *strongly agree*) with statements concerning whether or not gender bias was present in the encounter (e.g., "This situation is an example of gender discrimination"), and statements concerning their attitudes and intentions in response to the encounter (e.g., "I would tell Mike he is behaving unfairly"). Items for each set of vignettes were averaged to create a bias identification index (α = .78) and a bias false identification index (α = .83).

Growth mindset. Growth (vs. fixed) mindset was assessed using Neel and Shapiro's (2012) Lay Theories of Racial Bias Scale, reworded to refer to gender bias. Participants responded to three statements on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*), including, "People have a certain amount of gender bias and can't do much to change it" (reverse-coded). Items were averaged to create the growth mindset index (α = .90).

Self-efficacy. Self-efficacy was assessed using van Zomeren, Saguy, and Schellhaas's (2012) Individual Self-Efficacy Scale, adapted to refer specifically to gender bias in the workplace. Participants responded to three statements on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*), including, "I believe that I, as an individual, can help stop gender bias in the workplace." Items were averaged to create the self-efficacy index (α = .79).

Results

Bivariate correlations between all variables in Experiment 1 are presented in Table 2. All results from Experiment 1 are presented in Figure 1.

Awareness of bias. Consistent with previous tests of the VIDS intervention (Moss-Racusin et al., in press; Pietri et al., 2017), we observed a main effect of video on awareness of bias against women in the workplace, $F(1, 339) = 9.41$, $MSE = 0.67$,

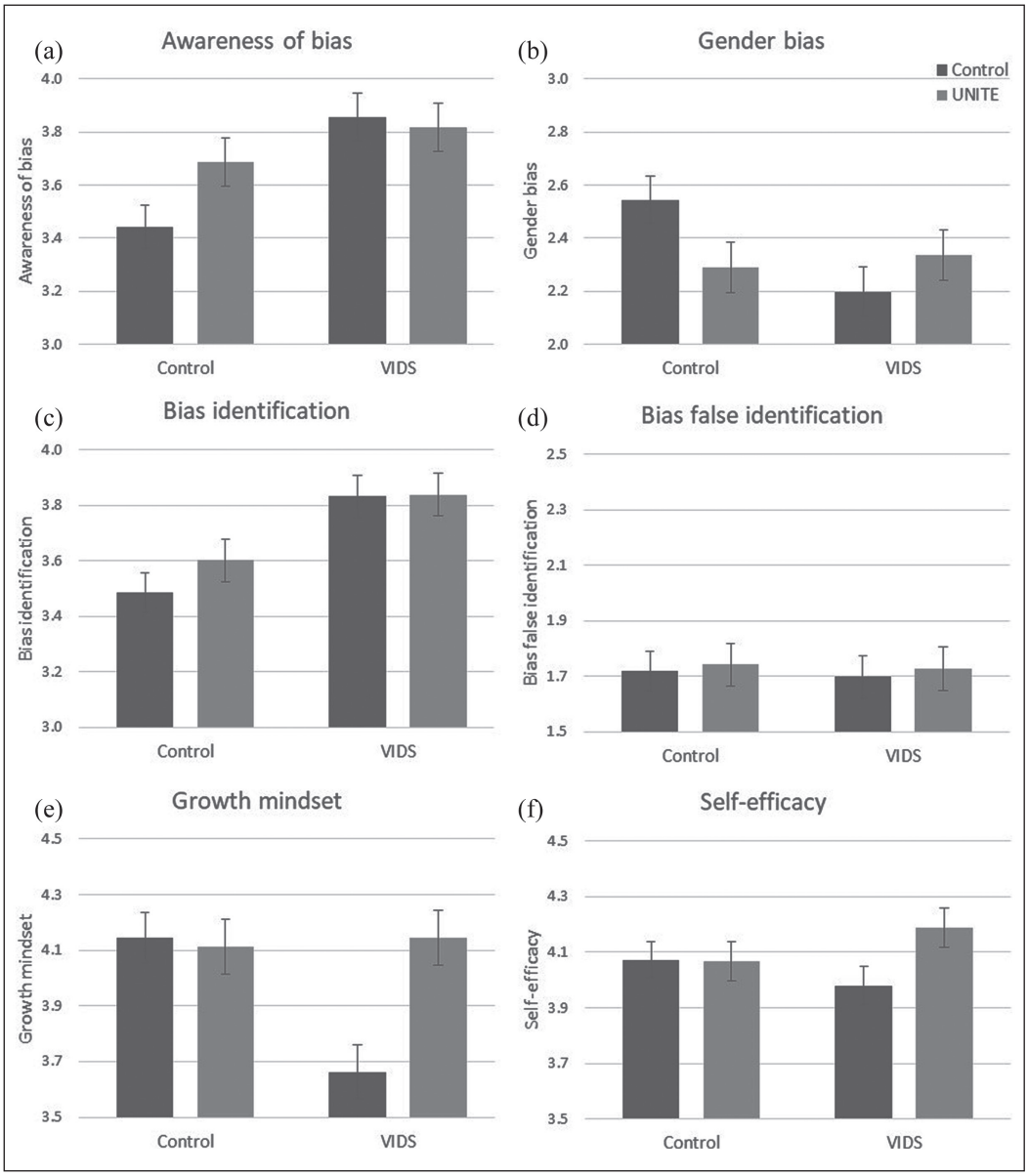

Figure 1. Bias outcomes among the general population (Experiment 1).
Note. Postintervention awareness of gender bias (a), gender bias (b), bias identification (d), bias false identification (d), growth mindset (e), and self-efficacy (f), by condition. Error bars indicate standard errors.

$p = .002$, $\eta^2 = .03$. There was no main effect of module, $F(1, 339) = 1.35$, $MSE = 0.67$, $p = .247$, $\eta^2 = .004$, nor a Video x Module interaction, $F(1, 339) = 2.63$, $MSE = 0.67$, $p = .106$, $\eta^2 = .01$.[2] As expected, participants in the VIDS+control condition reported significantly greater awareness of bias ($M = 3.86$, $SD = 0.76$) compared to those in the control+control condition ($M = 3.44$, $SD = 0.97$), $t(339) = 3.41$, $p = .001$, $d = 0.50$.[3] The effect of VIDS was not significant in the UNITE

module condition, $t(339) < 1$, $p = .747$. This appears to be due to a significant effect of UNITE among participants in the control video condition ($M = 3.69$, $SD = 0.77$), $t(339) = 1.99$, $p = .047$, $d = 0.30$, suggesting that UNITE alone also improves awareness of bias. However, there was no effect of module among participants in the VIDS video condition, $t(339) = -0.32$, $p = .747$, suggesting that UNITE does not further enhance VIDS' effectiveness in promoting awareness of bias. Together, these findings indicate that VIDS alone is sufficient to improve awareness of bias.

Gender bias. In regard to gender bias, the main effect of video condition was not statistically significant, $F(1, 339) = 2.61$, $MSE = 0.74$, $p = .107$, $\eta^2 = .01$, nor was the main effect of module, $F(1, 339) = 0.40$, $MSE = 0.74$, $p = .527$, $\eta^2 = .001$, but there was a significant Video x Module interaction, $F(1, 339) = 3.30$, $MSE = 0.74$, $p = .035$, $\eta^2 = .01$. In particular, consistent with previous research (Moss-Racusin et al., in press) and the pattern of results for awareness of bias, we observed that participants in the VIDS+control condition reported significantly lower modern sexism ($M = 2.20$, $SD = 0.76$) than did those in the control+control condition ($M = 2.54$, $SD = 1.02$), $t(339) = -2.71$, $p = .007$, $d = 0.40$. Those in the UNITE condition did not differ based on VIDS condition, $t(339) < 1$, $p = .732$. There was a significant effect of module among participants in the control video condition ($M = 2.29$, $SD = 0.80$), $t(339) = -1.97$, $p = .050$, $d = 0.30$, suggesting that UNITE alone also reduces gender bias. However, there was no effect of module among participants in the VIDS video condition, $t(339) = 1.02$, $p = .303$, suggesting that UNITE does not further enhance VIDS' effectiveness in reducing sexism. Together, these findings indicate that VIDS alone is sufficient to reduce gender bias.

Bias identification. We again observed a similar effect of VIDS in increasing participants' ability to correctly identify bias against women when presented in vignettes, replicating findings from Pietri et al. (2017). Overall, participants responded somewhat above the midpoint ($M = 3.68$, $SD = 0.78$) on a 5-point scale (1 = *strongly disagree*, 5 = *strongly agree*). Notably, there was a significant main effect of video, $F(1, 339) = 7.24$, $MSE = 0.48$, $p < .001$, $\eta^2 = .04$, but no main effect of module, $F(1, 339) = 0.66$, $MSE = 0.48$, $p = .416$, $\eta^2 = .002$, nor a Video x Module interaction, $F(1, 339) = 0.52$, $MSE = 0.48$, $p = .471$, $\eta^2 = .002$. As expected, participants in the VIDS+control condition were significantly more skilled at identifying bias ($M = 3.83$, $SD = 0.67$) than were those in the control+control condition ($M = 3.49$, $SD = 0.70$), $t(339) = 3.36$, $p = .001$, $d = 0.49$. VIDS also improved bias recognition for participants in the UNITE module condition, $t(339) = 2.17$, $p = .030$. In this case, UNITE alone did not improve bias identification, $t(339) = 1.10$, $p = .271$. Again, there was no effect of module among participants in the VIDS video condition, $t(339) = -0.06$, $p = .949$, suggesting that UNITE does not enhance VIDS' effectiveness in improving bias identification. Together, these findings indicate that VIDS alone is sufficient to improve bias identification.

Bias false identification. To examine whether VIDS might lead to overgeneralization of bias, we also examined participants' responses to the vignettes in which bias was absent. Overall, participants across conditions tended not to report bias in scenarios in which it was absent ($M = 1.72$, $SD = 1.03$). As observed by Pietri et al. (2017), no main or interactive effects were statistically significant, all Fs < 1, ps $> .74$. These findings importantly demonstrate that the intervention does not act to increase participants' tendency to claim bias indiscriminately.

Growth mindset. New to the current research, we observed a very different pattern of effects for growth mindset. In this case, there was a main effect of video, $F(1, 339) = 5.47$, $MSE = 0.79$, $p = .020$, $\eta^2 = .02$, a main effect of module, $F(1, 339) = 5.38$, $MSE = 0.79$, $p = .021$, $\eta^2 = .02$, and a Video x Module interaction, $F(1, 339) = 7.18$, $MSE = 0.79$, $p = .008$, $\eta^2 = .02$. As expected, participants in the VIDS+control condition reported a significantly more *fixed* mindset

(M = 3.66, SD = 1.03) compared to those in the control+control condition (M = 4.15, SD = 0.91), $t(339) = -3.66$, $p < .001$, $d = 0.53$. This unintended detrimental effect of VIDS was buffered in the UNITE condition (M = 4.14, SD = 0.79), $t(339) = 0.24$, $p = .814$. Similarly, we found that among participants in the control video condition, there was no effect of module, $t(339) = -0.26$, $p = .796$, but among those who viewed VIDS, UNITE significantly improved growth mindset, $t(339) = 3.49$, $p = .001$, $d = 0.53$. These findings indicate that, in isolation, VIDS resulted in the unintended consequence of promoting a fixed mindset about gender bias. However, UNITE successfully buffered against this negative effect.

Self-efficacy. As predicted, results for self-efficacy were similar to those obtained for mindset. In this case, there were no significant omnibus effects: main effect of video, $F(1, 339) = 0.05$, $MSE = 0.40$, $p = .822$, $\eta^2 = .00$; main effect of module, $F(1, 339) = 2.17$, $MSE = 0.40$, $p = .142$, $\eta^2 = .01$; interaction, $F(1, 339) = 2.49$, $MSE = 0.40$, $p = .116$, $\eta^2 = .01$. However, consistent with the growth mindset findings, we observed that there was no effect of module among participants in the control video condition, $t(339) = -0.07$, $p = .941$, but UNITE significantly improved self-efficacy among participants who viewed VIDS (M = 4.19, SD = 0.60 vs. M = 3.98, SD = 0.65), $t(339) = 2.13$, $p = .034$, $d = 0.33$. The simple effect of VIDS was not significant in either the control or UNITE module conditions, ps > .21. Together, these findings indicate that UNITE can be a useful supplement to VIDS for increasing self-efficacy to combat gender bias in the workplace.

Exploratory analyses. Some previous research has found gender bias interventions to be more effective for men than women (e.g., Jackson, Hillard, & Schneider, 2014). Consistent with prior research (e.g., Pietri et al., 2017; Swim et al., 1995), men in the current experiment tended to have lower levels of awareness of bias, higher levels of gender bias, and were less likely to label incidents as biased compared to women (see Table 2). However, consistent with previous work on VIDS (Moss-Racusin et al., in press; Pietri et al., 2017), there was no moderating effect of gender on any outcome in the current research, ps > .22. These results indicate that VIDS+UNITE is effective for men and women.

Discussion

Findings from Experiment 1 provide initial support for each of our three hypotheses: First, we replicated previous research demonstrating that VIDS improves awareness of bias in the workplace (Pietri et al., 2017) and reduces modern sexism (Moss-Racusin et al., in press) among both men and women. Importantly, we also replicated findings indicating that VIDS increases participants' ability to correctly recognize examples of gender bias as described in brief vignettes, but does not increase individuals' tendency to claim bias where it does not exist (Pietri et al., 2017). However, UNITE did not further enhance VIDS' effectiveness compared to VIDS alone, suggesting that VIDS is sufficient for improving bias literacy (Hypothesis 1).

Importantly, and in support of Hypothesis 2, we found that VIDS' effectiveness did not generalize to self-efficacy or growth mindset. Indeed, VIDS unintentionally led to a significantly more *fixed* mindset about gender bias compared to the control Videos. However, consistent with Hypothesis 3, UNITE was successful in buffering this effect and significantly restoring growth mindset and self-efficacy to address bias among individuals exposed to VIDS. Thus, a combined VIDS+UNITE intervention may be most successful for improving attitudes without impairing efficacy to take action.

Experiment 2

In Experiment 2, we sought to replicate our findings among STEM faculty. Following Moss-Racusin et al. (in press), we also sought to examine whether these effects would persist over time. Thus, we adopted a longitudinal design, including a baseline measurement two days before the intervention, an immediate postintervention measurement, and a follow-up one week after the intervention, to examine both the immediate and longer term effects of

VIDS and UNITE. This design also allowed us to analyze participants' change from baseline, as well as whether baseline levels of bias moderated the effectiveness of the intervention.

Method

Participants and recruitment. We recruited participants for Experiment 2 in collaboration with the Summer Institute (SI) on Scientific Teaching (Wood & Handelsman, 2004). There are seven SIs throughout the country focusing on training faculty to create more engaging science classrooms.[4] All 268 academic scientists who were scheduled to attend an SI during summer 2015 were invited to take part in an experiment that, "looks at how individuals react to and remember information from videos." Of those, 149 (56%) completed the experiment. This response rate is consistent with recruitment rates from SIs described in Moss-Racusin et al. (in press). Of importance, attrition rates across the three measurement time points were quite low, and compared favorably to those frequently obtained in longitudinal research (Capaldi & Patterson, 1987). Specifically, 133 participants (89% of the original sample) took part in the Time 2 (postintervention) session, and 130 participants (87% of the original sample) completed the Time 3 (follow-up) session.[5] All three sessions were completed prior to the SI. Participants were compensated with a $25 Amazon.com gift certificate after completing the Time 1 session, another $25 Amazon.com gift certificate after completing Time 2, and a final $50 Amazon.com gift certificate after completing Time 3. This data collection effort was identical to that utilized by Moss-Racusin et al. (in press).

Faculty participants (68% women; 77% White; age: $M = 42.90$, $SD = 10.82$, range = 25–73) were from diverse institution types, with a plurality from high research activity (Research I) universities (47% Research I universities, 18% Research II universities, 25% liberal arts colleges, and 10% primarily teaching colleges). Participants were in various career stages, with the majority being tenured or tenure-track professors (5% department chairs, 12% full professors, 18% associate professors, 26% assistant professors, 18% lecturers, 11% postdoctoral fellows or graduate students, and 10% other), and had taught for approximately ten years on average (range 0–40 years). The majority of participants worked in the biological sciences, but other STEM departments were also represented (75% biological sciences, 8% chemistry, 5% biomedical sciences, 3% physics, 2% mathematics, 1% engineering, 1% psychology, 4% other). On a scale of 1 (*strongly liberal*) to 7 (*strongly conservative*), participants' average political orientation was 2.23 ($SD = 1.41$). Previous research using samples from this population have found that individuals who choose to participate in research studies such as these do not differ systematically from SI attendees who choose not to participate (Moss-Racusin et al., in press).

Materials

VIDS. As in Experiment 1, participants were randomly assigned to either VIDS or the control video condition. Participants in the VIDS condition viewed three narrative videos and three corresponding expert interviews (counterbalanced to view either the narrative or expert interview video first for each pair), whereas participants in the control video condition viewed six science documentaries.

UNITE. As in Experiment 1, participants were randomly assigned to view either UNITE or a control module, modified from Experiment 1 to more directly relate to STEM education (vs. the general workplace). In particular, UNITE was adjusted from "addressing gender bias in the workplace" to "addressing gender bias in the classroom." The control module was adjusted from "creating an engaging workplace" to "creating an engaging science classroom." In particular, the control module discussed the importance of engaging student participation when teaching, and described helpful activities that can be easily added to a course (e.g., small group discussion).

Outcome measures. We administered the same measures of awareness of gender bias, gender bias, mindset, and self-efficacy as in Experiment

Table 3. Baseline bivariate correlation matrix (Experiment 2).

		1	2	3	4	5	6	7	8	9	10	11
1.	Awareness of bias											
2.	Gender bias	−.50***										
3.	Growth mindset	.02	−.05									
4.	Self-efficacy	.19*	−.14†	−.01								
5.	R1 institution	−.04	−.15†	.02	.00							
6.	Pretenure (vs. tenured)	.00	.04	.03	−.06	.08						
7.	Nontenure (vs. tenured)	.15†	−.09	.01	.06	.16*	−.46***					
8.	% Women faculty	−.22**	.15†	−.08	−.06	−.39***	−.11	−.03				
9.	% Women in lab	.09	−.07	−.02	.03	−.11	−.17*	.25**	.07			
10.	Conservatism	−.25**	.44***	−.07	.12	−.18*	.06	−.03	.13	.00		
11.	Age	−.18*	.06	.01	.03	−.10	−.47***	−.15†	.12	−.01	.06	
12	Gender (W = 0, M = 1)	−.34***	.12	−.03	−.14†	.09	.07	−.23**	−.03	−.30***	.06	.09

Note: R1 refers to high research activity institution.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

1, with two exceptions. First, references to women "in the workplace" were changed to women "in STEM." Second, due to time constraints and because these findings had already been replicated across multiple studies, we did not administer the bias identification and false identification vignette measures. Descriptive and reliability statistics for all scales are presented in Table 1, and baseline bivariate correlations between variables in Experiment 2 are presented in Table 3.

Procedure. Participants were contacted via email and invited to take part in an online study. At Time 1 participants provided baseline measures of all dependent variables as well as demographic information. Participants were recontacted two days later at Time 2 and randomly assigned to view either VIDS or the control videos followed by either UNITE or the control module. Subsequently, participants completed the postintervention dependent measures. Participants were then contacted a third time one week after the intervention to complete the same dependent measures at Time 3 (follow-up) and were then fully debriefed.

Results

Omnibus regression results are presented in Table 4, means and standard errors are presented in Table 5, effect sizes appear in Table 6, and graphs of the findings are illustrated in Figure 2.[6]

Awareness of bias. Consistent with Experiment 1, there was a significant omnibus postintervention effect of VIDS on awareness of bias against women in STEM, $b = .53$, $SE = 0.19$, $t(125) = 2.81$, $p = .006$. In line with previous findings, VIDS improved awareness of bias relative to the control video among those in control module condition, $b = .34$, $SE = 0.11$, $t(125) = 3.17$, $p = .002$. There was also significant improvement relative to baseline at postintervention for participants who had viewed VIDS, VIDS+control: $b = .41$, $SE = 0.08$, $t(125) = 4.98$, $p < .001$; VIDS+UNITE: $b = .18$, $SE = 0.07$, $t(125) = 2.61$, $p = .010$. However, UNITE did not increase the effectiveness of VIDS, and actually unexpectedly *weakened* VIDS' effectiveness at postintervention relative to VIDS alone, $b = -.23$, $SE = 0.11$, $t(125) = -2.11$, $p = .037$.

Table 4. Multilevel regression models for change in outcome variables relative to baseline (Experiment 2).

	Awareness of gender bias in STEM		Gender bias		Growth mindset		Self-efficacy	
	b	*SE*	*b*	*SE*	*b*	*SE*	*b*	*SE*
Intercept	.09	0.12	.16†	0.09	−.20	0.17	.16	0.15
Baseline	−.25***	0.05	−.24***	0.05	−.55***	0.09	−.32***	0.06
Time	−.03	0.07	−.09	0.06	.12	0.09	−.08	0.09
VIDS	.53**	0.19	−.37**	0.14	−.05	0.25	.18	0.23
UNITE	.01	0.19	.08	0.14	.44†	0.25	−.06	0.23
VIDS*UNITE	−.41	0.26	−.01	0.20	−.07	0.36	.01	0.32
VIDS*Time	−.18†	0.11	.16†	0.08	.05	0.14	−.10	0.13
UNITE*Time	.03	0.11	−.06	0.08	−.27*	0.14	.24†	0.13
VIDS*UNITE*Time	.14	0.15	.00	0.12	.19	0.19	−.06	0.19
$s^2_{Intercept}$	.09	0.02	.06	0.01	.33	0.05	.11	0.02
$s^2_{ResidualTime1}$	.09	0.02	.05	0.01	.13	0.04	.14	0.03
$s^2_{ResidualTime2}$	.09	0.02	.06	0.01	.17	0.04	.13	0.03

Note. Conditions were dummy coded with the control video and control module as the reference conditions. Time was coded with postintervention = 0 and follow-up = 1. Baseline score was grand mean centered. All models estimated a random person intercept and a variance components residual covariance matrix. The unstandardized regression coefficient is indicated by *b*, and the standard error of the estimate is indicated by *SE*.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 5. Means and standard errors for outcome variables by condition (Experiment 2).

Mean change from baseline (standard errors)

Variable		Postintervention		Follow-up	
		Control	VIDS	Control	VIDS
Awareness of gender bias in STEM	Control	0.06 (0.07)	0.41 (0.08)***	0.04 (0.07)	0.20 (0.08)*
	UNITE	0.10 (0.08)	0.18 (0.07)*	0.11 (0.08)	0.15 (0.07)*
Gender bias	Control	0.07 (0.05)	−0.13 (0.06)*	−0.02 (0.06)	−0.06 (0.07)
	UNITE	0.09 (0.06)	−0.12 (0.05)*	−0.05 (0.07)	−0.11 (0.06) †
Growth mindset	Control	−0.08 (0.11)	−0.08 (0.13)	0.04 (0.12)	0.09 (0.13)
	UNITE	0.09 (0.13)	0.21 (0.11) †	−0.07 (0.13)	0.29 (0.12)*
Self-efficacy	Control	0.08 (0.08)	0.16 (0.10) †	0.00 (0.08)	−0.01 (0.09)
	UNITE	0.26 (0.10)**	0.30 (0.08)***	0.41 (0.09)***	0.30 (0.08)***

Note. Least square mean estimates. Change from baseline significance tests.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

These patterns are generally similar at follow-up. No simple effects were statistically significant, *p*s > .14. However, participants in both the VIDS+control and VIDS+UNITE conditions continued to report significantly more awareness of bias relative to their baseline levels, $b = .20$, $SE = 0.08$, $t(125) = 2.47$, $p = .015$, and $b = .15$, $SE = 0.07$, $t(125) = 2.12$, $p = .036$, respectively. Taken together, these findings indicate that VIDS alone is sufficient for improving awareness of bias against women in STEM, and that these effects are still observable one week following intervention.

Table 6. Effect sizes (Experiment 2).

d effect sizes (compared to baseline)

Variable		Postintervention		Follow-up	
		Control	VIDS	Control	VIDS
Awareness of gender bias in STEM	Control	.14	.85	.06	.30
	UNITE	.22	.38	.17	.22
Gender bias	Control	.00	.36	.04	.15
	UNITE	.00	.34	.14	.30
Growth mindset	Control	.00	.00	.05	.11
	UNITE	.11	.26	.00	.38
Self-efficacy	Control	.14	.29	.00	.00
	UNITE	.46	.53	.76	.55

Note. Effect sizes were estimated by dividing the condition mean of the change score by the standard deviation of the change score at that time point. Effect sizes (*d*) indicate the number of standard deviations separating the condition mean at that time point from the condition mean at baseline. $d = 0.2$ is considered a small effect size, $d = 0.5$ is considered a medium effect size, and $d = 0.8$ is considered a large effect size (Cohen, 1988).

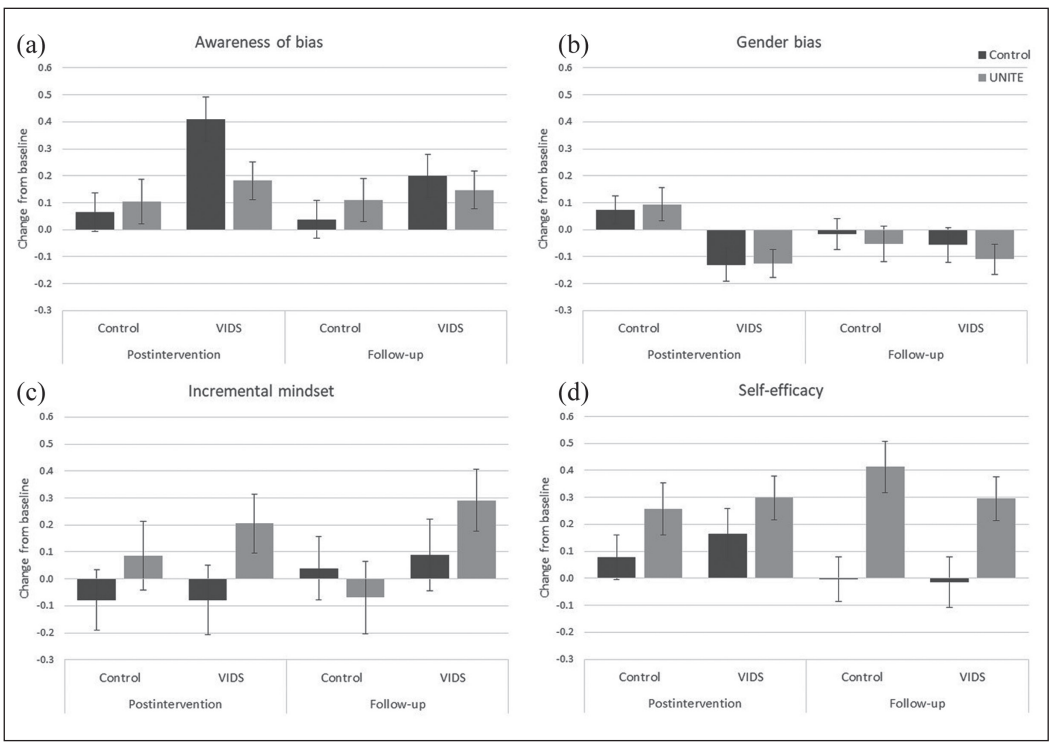

Figure 2. Change in bias outcomes among STEM faculty (Experiment 2).
Note. Change in (a) awareness of gender bias, (b) gender bias, (c) growth mindset, and (d) self-efficacy, relative to baseline by condition. Error bars indicate standard errors.

Gender bias. Similar results were observed for modern sexism. There was a significant omnibus effect of VIDS postintervention, $b = -.37$, $SE = 0.14$, $t(125) = -2.62$, $p = .010$. As expected, VIDS reduced gender bias relative to the control video both among those in the UNITE, $b = -.22$, $SE = 0.08$, $t(125) = -2.70$, $p = .008$, and the control module conditions, $b = -.20$, $SE = 0.08$, $t(125) = -2.50$, $p = .014$. There was also significant improvement relative to baseline at postintervention for participants who had viewed VIDS, VIDS+control: $b = -.13$, $SE = 0.06$, $t(125) = -2.11$, $p = .037$; VIDS+UNITE: $b = -.12$, $SE = 0.05$, $t(125) = -2.39$, $p = .019$. Consistent with Experiment 1, UNITE did not increase the effectiveness of VIDS. There was no significant simple effect of module in either the VIDS or the control video condition, ps > .80.

Patterns were somewhat similar at follow-up. No simple effects were statistically significant, ps > .24, although this appears to be driven by an unexpected but nonsignificant reduction in gender bias among participants in the two control video conditions. VIDS+UNITE is the only condition that remains different from baseline, $b = -.11$, $SE = 0.06$, $t(125) = -1.94$, $p = .055$, although the difference between VIDS+UNITE and VIDS+control is not statistically significant, $p = .543$. Although speculative and not hypothesized, this suggests that UNITE may have some utility in scaffolding the longer term effects of VIDS. Taken together, these findings indicate that VIDS alone is sufficient for reducing gender bias.

Growth mindset. Consistent with Experiment 1, the pattern of results for growth mindset was very different than the effects on bias awareness and gender bias. In particular, the omnibus effect of UNITE was marginally significant, $b = .44$, $SE = 0.25$, $t(125) = 1.73$, $p = .086$. Similar to Experiment 1, UNITE marginally improved growth mindset relative to the control module among those in the VIDS condition at postintervention, $b = .28$, $SE = 0.17$, $t(125) = 1.68$, $p = .096$, but did not impact upon participants in the control video condition, $b = .16$, $SE = 0.17$, $t(125) = 0.97$, $p = .335$. There was also marginally significant improvement relative to baseline at postintervention for participants who had viewed VIDS+UNITE, $b = .21$, $SE = 0.11$, $t(125) = 1.88$, $p = .063$, but not among participants who had viewed VIDS alone, $b = .08$, $SE = 0.13$, $t(125) = 0.61$, $p = .543$.

Interestingly, these effects strengthen over time. At follow-up, there was a statistically significant effect of VIDS among participants in the UNITE module condition, $b = .36$, $SE = 0.18$, $t(125) = 2.03$, $p = .043$. VIDS+UNITE also showed significant improvement from baseline, $b = .29$, $SE = 0.12$, $t(125) = 2.52$, $p = .013$, but, as predicted, VIDS+control did not, $b = .09$, $SE = 0.13$, $t(125) = 0.66$, $p = .511$. Taken together, these findings indicate that VIDS alone is not sufficient to increase growth mindset. UNITE can be a useful supplement to VIDS for improving the mindset that gender bias can be changed.

Self-efficacy. Finally, effects on self-efficacy were similar to those observed for mindset. As in Experiment 1, there were no significant omnibus effects at postintervention. However, we observed that both UNITE alone, $b = .26$, $SE = 0.10$, $t(125) = 2.68$, $p = .008$, and VIDS+UNITE led to improved self-efficacy relative to baseline, $b = .30$, $SE = 0.08$, $t(125) = 3.67$, $p < .001$ (while the control module conditions did not, ps > .09).

Consistent with the findings for growth mindset, these effects strengthen over time. At follow-up, there was a statistically significant effect of UNITE among participants in both the control video condition, $b = .42$, $SE = 0.13$, $t(125) = 3.30$, $p = .001$, and participants in the VIDS video condition, $b = .31$, $SE = 0.12$, $t(125) = 2.50$, $p = .014$. Both of these effects remain significantly different from baseline, control+UNITE: $b = .41$, $SE = 0.09$, $t(125) = 4.38$, $p < .001$; VIDS+UNITE: $b = .30$, $SE = 0.08$, $t(125) = 3.64$, $p < .001$. Together, these findings indicate that UNITE can be a useful supplement to VIDS for increasing self-efficacy to combat gender bias in the workplace.

Exploratory analyses. We also explored whether UNITE might have additional benefits for growth mindset and self-efficacy over and above merely buffering negative effects of VIDS. To do

so, we estimated the pairwise comparison between participants in the VIDS+UNITE and the control video+control module condition. We found that participants in the VIDS+UNITE condition reported marginally more growth mindset than participants in the control+control condition postintervention, $b = .28$, $SE = 0.16$, $t(125) = 1.81$, $p = .072$, $d = 0.36$, and somewhat more at follow-up, $b = .25$, $SE = 0.17$, $t(125) = 1.53$, $p = .129$, $d = 0.33$. VIDS+UNITE participants also reported marginally more self-efficacy than did participants in the control+control condition postintervention, $b = .22$, $SE = 0.12$, $t(125) = 1.89$, $p = .061$, $d = 0.39$, and significantly more at follow-up, $b = .30$, $SE = 0.12$, $t(125) = 1.53$, $p = .011$, $d = 0.56$. This indicates that UNITE tended to not only restore self-efficacy and growth mindset among VIDS participants, but also heighten them relative to a control condition.

Consistent with Experiment 1, men reported lower awareness of bias than did women (Table 3). However, in this case the difference in gender bias itself was not statistically significant. Moreover, there was again no moderating effect of participant gender on any outcome, ps $> .12$, indicating that VIDS+UNITE is generally equally effective for men and women. There were also no moderating effects of baseline scores for any of the four outcomes with one exception: for awareness of bias, individuals who were less aware of bias at baseline were more positively affected by both VIDS, $b = .79$, $SE = 0.31$, $t(121) = 2.57$, $p = .011$, and UNITE, $b = .62$, $SE = 0.31$, $t(121) = 1.99$, $p = .049$. Notably, this effect was particularly apparent among men, Awareness of Bias x Gender x VIDS x UNITE: $b = -.24$, $SE = 0.91$, $t(113) = -2.62$, $p = .010$. This suggests that the intervention may be especially effective for individuals, and particularly men, who had more negative baseline attitudes.[7] However, this moderation was not observed on the other three outcomes, nor in previous research on STEM faculty (Moss-Racusin et al., in press), so future research is necessary to determine if VIDS+UNITE might have an especially positive impact for individuals with lower a priori awareness of gender bias in STEM.

Discussion

Experiment 2 provided additional support for all three hypotheses. Supporting Hypothesis 1, VIDS improved awareness of bias against women in STEM both immediately postintervention and one week later. UNITE did not increase the effectiveness of VIDS. Similarly, VIDS led to a reduction in gender bias postintervention, although this effect weakened somewhat one week later. In contrast, and consistent with Hypothesis 2, we again observed that VIDS alone was ineffective in increasing self-efficacy or a growth mindset. Supporting Hypothesis 3, we found that supplementing VIDS with UNITE led to significant improvements in growth mindset and self-efficacy relative to baseline, and these effects persisted over time. In sum, these findings provide additional evidence that a combined VIDS+UNITE intervention may be most successful for both improving attitudes and increasing efficacy to take action.

General Discussion

Lingering gender biases continue to contribute to the underrepresentation of women in STEM fields (e.g., Milkman, Akinola, & Chugh, 2012, 2015; Moss-Racusin et al., 2012; Reuben, Sapienza, & Zingales, 2014). Although evidence-based gender bias interventions have been found to be effective in randomized controlled trials, they are generally resource-intensive and difficult to administer widely (Carnes et al., 2015; Zawadzki et al., 2012). Recently, VIDS has been developed as an easily administrable and scalable alternative that effectively improves awareness of gender bias, identification of bias, and gender bias itself among both the general population and STEM faculty (Moss-Racusin et al., in press; Pietri et al., 2017). However, previous trials of VIDS have not assessed VIDS' impact on self-efficacy to combat bias or perceptions that gender bias can be overcome, and evidence from the health and organizational behavior literatures suggests that providing individuals with information about the pervasiveness of bias without

offering actionable tools for addressing the problem can lead individuals to feel that the problem is socially normative and insurmountable (Bandura, 2004; Duguid & Thomas-Hunt, 2015; Floyd et al., 2000; Good & Abraham, 2011). To address this issue, we supplemented VIDS with an educational module that encouraged a growth mindset about the ability to change gender bias and provided concrete, empirically sound instructions for promoting equity and addressing implicit biases.

As in previous research, we found that VIDS increased awareness of bias and reduced sexism. VIDS also helped increase participants' ability to identify bias while maintaining their ability to differentiate bias from fair treatment (cf. Ceci et al., 2014; Haidt, 2017). However, consistent with past research on the effects of exposure to negative information (e.g., Carr et al., 2012), VIDS was ineffective at supporting a growth mindset and did not lead to self-efficacy to address bias. The newly developed UNITE module buffered against perceptions that gender bias is unchangeable and restored participants' self-efficacy to address bias among both the general population and STEM faculty.

Together, these results suggest that interventions that merely aim to improve attitudes or increase awareness may have unintended negative consequences of leading intervention participants to feel inefficacious to take action. However, interventions that additionally buffer against fixed mindsets about bias and provide individuals with tools to confront it may be more likely to result in equitable and inclusive workplaces. Thus, combining multiple effective intervention components may result in a particularly powerful and successful diversity training. More broadly, the current findings demonstrate that researchers should continue to dynamically refine even powerful interventions such as VIDS to ensure that they are maximally effective and circumvent unintended negative consequences.

In that spirit, we suggest several directions for future research. Consistent with prior VIDS research (Moss-Racusin et al., in press; Pietri et al., 2017), we found VIDS+UNITE to be successful for men and women and those high and low on baseline prejudice. Even women and individuals with greater a priori bias literacy tended to have "room to grow." This also indicates that VIDS+UNITE can be administered indiscriminately to entire organizations with the expectation of benefit to the group as a whole. However, there was some evidence that those lower on awareness of bias (and particularly men) were more significantly impacted by the intervention. Indeed, practitioners may be particularly interested in interventions that work to "close the gap" between women and men (or those more and less bias literate). Future research should continue to investigate the impact of VIDS and other interventions on low bias literate subgroups.

Second, it is promising that the UNITE module, which consisted merely of a silent PowerPoint presentation, led to improvements in growth mindset and self-efficacy. Nevertheless, the effect sizes were somewhat small. Moreover, given the small sample sizes, particularly in Experiment 2, the precise impact of VIDS and UNITE on self-efficacy remains unclear. It is possible that a module that is more transporting would lead to greater responsiveness to the module and stronger effects (see also Pietri et al., 2017).

Third, we hypothesized that learning that bias is malleable would lead to improvements in growth mindset and that learning concrete strategies for effecting change would lead to improvements in self-efficacy. However, it is possible that learning that bias is malleable could improve self-efficacy (or that learning concrete strategies could improve mindset). Future research should differentiate these possibilities.

Fourth, future research should aim to assess behavioral outcomes. The module improved two psychological constructs that previous research has found to be critical to effective behavioral action: growth mindset (e.g., Rattan & Dweck, 2010) and self-efficacy (Bandura, 1977; Carnes et al., 2015; Sevo & Chubin, 2010). However, whether the module will inspire actual behavior aimed at reducing gender discrimination remains to be empirically examined. Finally, although our aim in the present research was to reduce bias

against women, social identities are complex and intersectional, and it is unknown whether our intervention generalizes more broadly to the treatment of individuals with other, or multiple, marginalized identities. Future research is needed to ensure that *all* individuals, such as those who identify outside the gender binary or who are members of stigmatized racial or sexual orientation groups, have access to equitable and inclusive workplaces.

Conclusion

The current research demonstrates that bias interventions that encourage a growth mindset and give people the tools to promote equity are promising methods for improving gender equity in STEM. Such interventions can offer an extra boost to successful awareness of bias programs like VIDS and help ameliorate unintended negative consequences, ensuring that program participants feel efficacious to promote diversity. In the long run, STEM workplaces that are equitable and inclusive will ensure that the most talented individuals have the opportunity to contribute to scientific advancement.

Acknowledgements

This manuscript is partially based on a Science Research Independent Study at Jefferson High School, Lafayette, IN, by the fourth author under the supervision of the first author, presented at the Lafayette Regional Science and Engineering Fair. The authors thank the leaders of the Summer Institutes for assistance with participant recruitment, playwright Dipika Guha, biological sciences consultants Matthew Akamatsu and Jessica Miles, Sean P. Lane for advice about data analysis, Zachary Chacko for assistance with database management, and members of the Social Cognition of Social Justice Lab, Michelle Ryan, and two anonymous reviewers for helpful feedback on an earlier draft of this paper.

Funding

This research was funded in part by Alfred P. Sloan Foundation Grants #213-3-15 to the third and last author and #B2013-38 to the third author, and a Howard Hughes Medical Institute Professor grant to the last author.

Notes

1. Because participants who do not work outside the home may have unrealistically high expectations of their hypothetical likelihood of confronting gender bias in the workplace (Woodzicka & LaFrance, 2001), we elected to exclude their responses from analysis as a more conservative test of our hypotheses. However, work status did not moderate any of our effects, and the statistical significance of all analyses of variance remained consistent among the full sample.
2. Our specific hypotheses center on (a) replicating the simple effect of VIDS in the control module condition on awareness of bias, gender bias, and identification (but not false identification) of bias reported by Pietri et al. (2017); and (b) examining the simple effects of UNITE in the VIDS and control video conditions on growth mindset and self-efficacy. Because we hypothesize different patterns of effects depending on the outcome variable, for the sake of clarity and completeness we report both the omnibus and simple effects analyses for all six dependent variables (see also, Games, 1973; Howell, 2013; Wilcox, 1987).
3. All *d* effect sizes were constructed using the overall sample *SD*.
4. To estimate 80% power to observe significant effects of VIDS+UNITE, we conducted a power analysis in SAS Version 9.4 (Lane & Hennes, 2018). We simulated data using two sources of information. Because the effect of VIDS on awareness of bias and gender bias was a direct replication of Moss-Racusin et al. (in press), we estimated power to detect these postintervention effects using the hybrid condition effects reported in their Table 6. To estimate power for growth mindset and self-efficacy, which had not been explored previously, we used the effects reported in Experiment 1 of the current manuscript. Regarding the latter two outcomes, we estimated effects to be the same at both postintervention and follow-up, given previous evidence of limited decay of VIDS' effectiveness (Moss-Racusin et al., in press). Additionally, Moss-Racusin et al. (in press) reported the residual variance across outcomes among their faculty sample to be approximately 10–20% of the residual among MTurk samples, and the random intercept to be approximately the same magnitude as the residual. Therefore, we estimated the random intercept and both the postintervention

and follow-up residuals to be 20% of the mean square error reported in Experiment 1 of the current research. These simulations determined that we would need 50 participants to power main effects of VIDS on awareness of bias, 88 participants to power main effects of VIDS on modern sexism, and 150 participants to power the omnibus VIDS+UNITE interaction effect on growth mindset at Time 2. The simulations revealed that successful recruitment of all 268 academic scientists participating in the SI would still be insufficient to reliably observe significant omnibus effects of VIDS+UNITE on self-efficacy. Nevertheless, because of the unique opportunity to explore the current research questions using participants from our primary population of interest, who tend to be more difficult to recruit than members of the general population, we decided to retain measures of self-efficacy in Experiment 2. However, we anticipated results for this outcome to likely be imprecise and statistically nonsignificant.

5. Dropout was not significantly predicted by experimental condition (ps > .92) nor by baseline levels of any dependent variable (ps > .093).
6. Because our design was longitudinal, we analyzed the data using multilevel modeling (Raudenbush & Bryk, 2002) with the mixed procedure in SAS Version 9.4. Identical to Moss-Racusin et al. (in press), we adopted regressed change models in which change from baseline was predicted by video condition, module condition, time, and their interactions, adjusting for baseline score (McArdle, 2009). We estimated a random person intercept and allowed the residual variation at Times 2 and 3 to be freely estimated but constrained the residual covariance to be 0 (i.e., a variance components matrix). We initially allowed the residual covariance to be freely estimated, but the estimate was negligible and prohibited model convergence.
7. Although VIDS participants tend not to differ on any demographic characteristic from nonparticipants who attend an SI (Moss-Racusin et al., in press), it remains unknown whether they differ from STEM scientists who do not attend SIs. Indeed, the majority of the participants were biologists, a field that has historically had greater representation of women than many other STEM disciplines (NSF, 2017). This might raise the concern that VIDS participants may be especially receptive to VIDS. However, the finding that those *lowest* on awareness of gender bias were most responsive to the intervention provides some evidence that these effects are likely to replicate more broadly across other STEM fields.

ORCID iD

Corinne A. Moss-Racusin https://orcid.org/0000-0002-9824-7524

References

Bandura, A. (1977). *Self-efficacy: The exercise of control*. New York, NY: Worth.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior*, *31*, 143–164. doi:10.1177/1090198104263660

Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin*, *142*, 1227–1274. doi:10.1037/bul0000067

Bilimoria, D., & Liang, X. (2013). State of knowledge about the workforce participation, equity, and inclusion of women in academic science and engineering. In M. Wyer, M. Barbercheck, D. Cookmeyer, H. O. Ozturk & M. Wayner (Eds.), *Women, science and technology: A reader in feminist science studies* (pp. 21–50). London, UK: Routledge.

Block, L. G., & Keller, P. A. (1995). When to accentuate the negative: The effects of perceived efficacy and message framing on intentions to perform a health-related behavior. *Journal of Marketing Research*, *32*, 192–203. doi:10.2307/3152047

Capaldi, D., & Patterson, G. R. (1987). An approach to the problem of recruitment and retention rates for longitudinal research. *Behavioral Assessment*, *9*, 169–177.

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*, 63–77. doi:10.1037/a0028128

Carnes, M., Devine, P. G., Manwell, L. B., Byars-Winston, A., Fine, E., Ford, C. E., . . . Sheridan, J. (2015). The effect of an intervention to break the gender bias habit for faculty at one institution: A cluster randomized, controlled trial. *Academic Medicine*, *90*, 221–230. doi:10.1097/ACM.0000000000000552

Carr, P. B., Dweck, C. S., & Pauker, K. (2012). "Prejudiced" behavior without prejudice? Beliefs about the malleability of prejudice affect interracial interactions. *Journal of Personality and Social Psychology*, *103*, 452–471. doi:10.1037/a0028849

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, *15*, 75–141. doi:10.1177/1529100614541236

Ceci, S. J., & Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences*, *108*, 3157–3162. doi:10.1073/pnas.1014871108

Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin*, *25*, 413–423. doi:10.1177/0146167299025004002

Cheryan, S. (2012). Understanding the paradox in math-related fields: Why do some gender gaps remain while others do not? *Sex Roles*, *66*, 184–190. doi:10.1007/s11199–011–0060-z

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*, 1045–1060. doi:10.1037/a0016239

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cortina, L. M. (2008). Unseen injustice: Incivility as modern discrimination in organizations. *Academy of Management Review*, *33*, 55–75. doi:10.5465/AMR.2008.27745097

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating the sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, *32*, 917–929. doi:10.1177/0146167206287721

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, *21*, 1051–1057. doi:10.1177/0956797610377342

Dobbin, F., & Kalev, A. (2016). Why diversity programs fail. *Harvard Business Review*, *94*, 52–61.

Duguid, M. M., & Thomas-Hunt, M. C. (2015). Condoning stereotyping? How awareness of stereotyping prevalence impacts expression of stereotypes. *Journal of Applied Psychology*, *100*, 343–359. doi:10.1037/a0037908

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

Floyd, D. L., Prentice-Dunn, S., & Rogers, R. W. (2000). A meta-analysis of research on protection motivation theory. *Journal of Applied Social Psychology*, *30*, 407–429. doi:10.1111/j.1559–1816.2000.tb02323.x

Games, P. A. (1973). Type IV errors revisited. *Psychological Bulletin*, *80*, 304–307. doi:10.1037/h0034832

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist*, *56*, 109–118. doi:10.1037/0003–066X.56.2.109

Good, A., & Abraham, C. (2011). Can the effectiveness of health promotion campaigns be improved using self-efficacy and self-affirmation interventions?: An analysis of sun protection messages. *Psychology & Health*, *26*, 799–818. doi:10.1080/08870446.2010.495157

Haidt, J. (2017). The unwisest idea on campus: Commentary on Lilienfeld (2017). *Perspectives on Psychological Science*, *12*, 176–177. doi:1745691616667050

Howell, D. C. (2013). *Statistical methods for psychology* (8th ed.). Belmont, CA: Wadsworth.

Jackson, S. M., Hillard, A. L., & Schneider, T. R. (2014). Using implicit bias training to improve attitudes toward women in STEM. *Social Psychology of Education*, *17*, 419–438. doi:10.1007/s11218–014–9259–5

Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationships research. *Journal of Social and Personal Relationships*, *35*, 7–31. doi:10.1177/0265407517710342

Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of anti-prejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, *22*, 1472–1477. doi:10.1177/0956797611427918

Liben, L. S., & Coyle, E. F. (2014). Developmental interventions to address the STEM gender gap: Exploring intended and unintended consequences. *Advances in Child Development and Behavior*, *47*, 77–115. doi:10.1016/bs.acdb.2014.06.001

Lindberg, S. M., Hyde, J., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, *136*, 1123–1135. doi:10.1037/a0021276

McArdle, J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577–605. doi:10.1146/annurev.psych.60.110707.163612

Milkman, K. L., Akinola, M., & Chugh, D. (2012). Temporal distance and discrimination: An audit study in academia. *Psychological Science*, *23*, 710–717. doi:10.1177/0956797611434539

Milkman, K. L., Akinola, M., & Chugh, D. (2015). What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations. *Journal of Applied Psychology*, *100*, 1678–1712. doi:10.1037/apl0000022

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*, 16474–16479. doi:10.1073/pnas.1211286109

Moss-Racusin, C. A., Pietri, E., Hennes, E. P., Dovidio, J. F., Brescoll, V. L., Roussos, G., & Handelsman, J. (in press). Reducing STEM gender bias with VIDS (Video interventions for diversity in STEM). *Journal of Experimental Psychology: Applied*.

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. doi:10.1037/a0018093

National Science Foundation. (2017). *Women, minorities, and persons with disabilities in science and engineering: 2017* (Special Report NSF 17–310). Retrieved from www.nsf.gov/statistics/wmpd

Neel, R., & Shapiro, J. R. (2012). Is racial bias malleable? Whites' lay theories of racial bias predict divergent strategies for interracial interactions. *Journal of Personality and Social Psychology*, *103*, 101–120. doi:10.1037/a0028237

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44–59. doi:10.1037/0022–3514.83.1.44

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., . . . Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*, 36–88. doi:10.1080/10463280701489053

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, *96*, 574–587. doi:10.1037/a0011989

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. doi:10.1146/annurev.psych.60.110707.163607

Penner, A. M. (2015). Gender inequality in science. *Science*, *347*, 234–235. doi:10.1126/science.aaa3781

Pietri, E. S., Moss-Racusin, C. A., Dovidio, J. F., Guha, D., Roussos, G., Brescoll, V. L., & Handelsman, J. (2017). Using Video to increase gender bias literacy toward women in science. *Psychology of Women Quarterly*, *41*, 175–196. doi:10.1177/0361684316674721

Rattan, A., & Dweck, C. (2010). Who confronts prejudice? The role of implicit theories in the motivation to confront prejudice. *Psychological Science*, *21*, 952–959. doi:10.1177/0956797610374740

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: SAGE.

Renzulli, L. A., Grant, L., & Kathuria, S. (2006). Race, gender, and the wage gap: Comparing faculty salaries in predominately White and historically Black colleges and universities. *Gender & Society*, *20*, 491–510. doi:10.1177/0891243206287130

Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, *111*, 4403–4408. doi:10.1073/pnas.1314788111

Rudman, L. A., & Glick, P. (2008). *The social psychology of gender*. New York, NY: Guilford.

SAS (Version 9.4) [Computer software]. Cary, NC: SAS Institute Inc.

Settles, I. H., Cortina, L. M., Buchanan, N. T., & Miner, K. (2013). Derogation, discrimination, and (dis)satisfaction with jobs in science: A gendered analysis. *Psychology of Women Quarterly*, *37*, 179–191. doi:10.1177/0361684312468727

Sevo, R., & Chubin, D. E. (2010). Bias literacy: A review of concepts in research on gender discrimination and the U.S. context. In A. Carter-Steel & E. Cater (Eds.), *Women in engineering: Science and technology: Education and career challenges* (pp. 21–49). Hershey, PA: IGI Global.

Sexism has no place in science [Editorial]. (2015). *Nature*, *522*, 255. doi:10.1038/522255a

Stephenson, M. T., & Witte, K. (1998). Fear, threat, and perceptions of efficacy from frightening skin cancer messages. *Public Health Reviews*, *26*, 147–174.

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology*, *68*, 199–214. doi:10.1037/0022–3514.68.2.199

Swim, J. K., Hyers, L. L., Cohen, L. L., & Ferguson, M. J. (2001). Everyday sexism: Evidence for its incidence, nature, and psychological impact from three daily diary studies. *Journal of Social Issues*, *57*, 31–53. doi:10.1111/0022–4537.00200

Van Zomeren, M., Saguy, T., & Schellhaas, F. M. H. (2012). Believing in "making a difference" to collective efforts: Participative efficacy beliefs as a unique predictor of collective action. *Group Processes & Intergroup Relations*, *16*, 618–634. doi:10.1177/1368430212467476

Wilcox, R. R. (1987). New designs in analysis of variance. *Annual Review of Psychology*, *38*, 29–60. doi:10.1146/annurev.ps.38.020187.000333

Williams, W. M., & Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*, *112*, 5360–5365. doi:10.1073/pnas.1418878112

Wood, W. B., & Handelsman, J. (2004). Meeting report: The 2004 National Academies Summer Institute on Undergraduate Education in Biology. *Cell Biology Education*, *3*, 215–217. doi:10.1187/cbe.04-07-0057

Woodzicka, J. A., & LaFrance, M. (2001). Real versus imagined gender harassment. *Journal of Social Issues*, *57*, 15–30. doi:10.1111/0022-4537.00199

Wright, A. L., Schwindt, L. A., Bassford, T. L., Reyna, V. F., Shisslak, C. M., Germain, P. A. S., & Reed, K. L. (2003). Gender differences in academic advancement: Patterns, causes, and potential solutions in one US college of medicine. *Academic Medicine*, *78*, 500–508.

Zawadzki, M. J., Danube, C. L., & Shields, S. A. (2012). How to talk about gender inequity in the workplace: Using WAGES as an experiential learning tool to reduce reactance and promote self-efficacy. *Sex Roles*, *67*, 605–616. doi:10.1007/s11199–012–0181