*This is an Accepted Manuscript of an article published by Taylor & Francis in The Journal of Higher Education on April 11, 2018, available online:*
*https://doi.org/10.1080/00221546.2018.1437665*

# Testing an Intervention for Recognizing and Reporting Subtle Gender Bias in Promotion and Tenure Decisions

Jessica L. Cundiff

Missouri University of Science and Technology

Cinnamon L. Danube and Matthew J. Zawadzki

University of California, Merced

Stephanie A. Shields

The Pennsylvania State University

Author Note

Correspondence concerning this manuscript should be addressed to Jessica L. Cundiff at cundiffjl@mst.edu or Cinnamon Danube at cdanube@merced.edu.

Abstract

Women make up the majority of doctoral degree earners, yet remain underrepresented in tenure-track positions within the academy. Gender disparities result in part from the accumulation of subtle, typically unintentional, biases that pervade workplace structures, practices, and patterns of interactions that inadvertently favor men. However, the subtle nature of gender bias makes it difficult to detect, thus diminishing the likelihood of action to address it. We experimentally evaluated the effectiveness of a brief intervention, the Workshop Activity for Gender Equity Simulation in the Academy (WAGES-Academic), which was designed to increase recognition of subtle gender bias in the academic workplace. Participants ($N$ = 177) completed either the WAGES intervention or one of two control conditions and later evaluated promotion and tenure materials of a woman faculty member who received either a blatant sexist, subtle sexist, or nonsexist review. Consistent with hypotheses, WAGES participants (vs. controls) detected more subtle gender bias ($p$s < .02) and were subsequently more likely to report concerns about bias ($p$s < .04). Results suggest that low-cost interventions that educate individuals about subtle bias in a non-threatening way may increase detection and reporting of gender bias in higher education institutions.

*Keywords:* active learning, experiential learning, discrimination, intervention, sexism, gender bias

# Testing an Intervention for Recognizing and Reporting Subtle Gender Bias in Promotion and Tenure Decisions

Women have made substantial progress in academia, holding a larger proportion of tenure-track faculty positions today than ever before (U.S. Department of Education, 2015). Yet, women continue to receive lower salaries than men and remain underrepresented in tenure-track positions and the upper ranks of faculty and administration, particularly in the physical sciences, technology, and engineering (Catalyst, 2015; Curtis, 2011). Scholars estimate that, barring any purposeful intervention, current trends will not reach gender equity for more than 100 years (Cress & Hart, 2009; Marschke, Laursen, Nielsen, & Rankin, 2007; Moss-Racusin et al., 2014).

Intervention is thus needed to reach gender equity in the academy within our lifetimes. Yet interventions can be time-consuming, costly and, if not administered appropriately, can backfire. To address these issues we examined the effectiveness of a brief, low-cost intervention – the Workshop Activity for Gender Equity Simulation in the Academy (WAGES-Academic, referred to hereafter as WAGES; Shields, Zawadzki, & Johnson, 2011; http://wages.la.psu.edu/) – that employs active learning strategies to educate about subtle gender bias in the academy. We report on a controlled experimental evaluation of this intervention tool for increasing the detection and reporting of subtle gender bias in promotion and tenure decisions.

## Subtle Gender Bias in the Academy

Multiple factors contribute to gender disparities in the academy. One factor that has received considerable empirical support is the role of subtle yet systemic gender biases that pervade workplace structures, practices, and patterns of interactions in ways that inadvertently favor men (Cooper et al., 2007; Valian, 1999) and that differentially affect women at different intersectional positions (Muhs, Niemann, González, & Harris, 2012). Subtle gender bias is

rooted in culturally-derived, broad assumptions about gender that ascribe greater value, competence, and leadership qualities to men than women (Heilman, 2012). These stereotypes bias evaluations of, and behavior toward, women and men, typically unconsciously and unintentionally.

One way that stereotypes bias evaluations is by influencing the standards by which work is evaluated. Stereotypic expectations about women's performance lead people to apply harsher standards when evaluating women than men (Heilman, 2012). In a study of medical postdoctoral fellowships, women applicants needed substantially more publications than men applicants to be seen as equally competent by reviewers (Wenneras & Wold, 1997). Similarly, in a controlled experimental study, university faculty rated identical job applicants as less hirable when the applicant was a woman compared to man, and were four times more likely to write cautionary comments (e.g., "I would need to see evidence that she had gotten these grants and publications on her own") for the woman than man applicant (Steinpreis, Andres, & Ritzke, 1999). A recent wide-scale experimental study revealed that these gender biases have not diminished: university faculty rated identical job candidates as less hirable, offered the candidate a lower salary, and offered less mentoring when the candidate was a woman compared to man (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012). Faculty are also less responsive to inquiries from prospective doctoral candidates who are women compared to men (Milkman, Akinola, & Chugh, 2015). These experimental studies echo findings from qualitative studies in which women faculty report experiencing harsher scrutiny and fewer rewards for comparable work compared to men faculty (Hart, 2016). Women faculty of color, who contend with racialized gender bias, are particularly likely to experience double standards and trivialization of their work (Muhs et al., 2012; Turner, Gonzalez, & Wong, 2011).

Gender stereotypes also contribute to gendered divisions of labor, such that stereotypically nurturing roles (e.g., teaching; service) are more often assigned to women than men (Hart, 2016; Hart & Cress, 2008; O'Meara, Kuvaeva, Nyunt, Jackson, & Waugaman, 2017). Women faculty of color are especially vulnerable to extra service work; the small proportion of women faculty of color on university campuses make them targets of frequent requests and expectations to serve as a voice for minority individuals (Hirshfield & Joseph, 2012). However, these types of service roles tend to be devalued by the academic merit system (Martinez Aleman, 2014).

Stereotypes also guide expectations about how women and men ought to behave. When people violate these expectations by acting outside their prescribed gender role, they risk facing social and economic repercussions (Rudman & Phelan, 2008). This creates a double bind for women because the traits and roles necessary for success in academia (e.g., assertiveness; self-promotion; research roles) are stereotypically associated with men, whereas the traits and roles that are devalued (e.g., supportive; nurturing; teaching and service roles) are stereotypically associated with women (Martinez Aleman, 2014). If women meet their gender role expectations by emphasizing teaching and service, they are less likely to advance. Yet, if women emphasize the masculine traits and roles necessary for advancement, they are penalized with harassment, marginalization, and challenges to credibility for diverging from gender role expectations (Cooper et al., 2007; Hart, 2016; Lester, 2011). The exact type of behavior that elicits punishment differs across racial groups (e.g., assertive behavior from white women; self-promoting behavior from black women; Livingston, 2013), but the notion is similar: white women and women of color must act like stereotypical white men to demonstrate competence,

but when they do, they are disliked and suffer social and economic repercussions that hinder their career advancement (Rudman & Phelan, 2008).

**The Need for Intervention**

Subtle gender bias does not typically result from a conscious desire to impede women's career progress. Rather, the crux of the problem is that decision making, in addition to deliberate evaluation of information, typically employs a range of cognitive shortcuts that unconsciously and automatically fill in missing information. Stereotypes are one type of cognitive shortcut that make it possible to process large quantities of complex information, but also make people susceptible to unintentional errors and biased judgments (Kahneman, 2003). Gender stereotypes are widespread and culturally ingrained such that they are automatically activated for most people (Nosek et al., 2007) and can unconsciously influence the behavior of even the most egalitarian individual (Devine, 1989), women and men alike (Moss-Racusin et al., 2012). The result is that, unless intentionally interrupted (Williams, 2014), unconscious gender stereotypes covertly interfere with interpersonal interactions and decision-making in ways that create a chilly climate and invisible barriers to women faculty's job satisfaction, productivity, and career advancement (Clark & Corcoran, 1986; Cress & Hart, 2009; Monroe, Ozyurt, Wrigley, & Alexander, 2008).

Not only is it difficult to control personal bias, but it is also difficult to detect bias when it occurs. Single incidents of bias, such as gender differences in salary increases, often lack appropriate comparative data and are based on subjective criteria that can be easily justified on the basis of some factor other than gender, making it difficult to identify as discriminatory (Crosby, Clayton, Alksnis, & Hemker, 1986; Heilman & Haynes, 2008). Even when single incidents are overtly discriminatory and easy to identify, such as making a stereotypic comment

about women's abilities, people fail to recognize that these minor incidents are harmful and part of a larger problem (Cundiff, Zawadzki, Danube, & Shields, 2014; Swim, Mallett, Russo-Devosa, & Stangor, 2005).

However, single minor incidents of bias accumulate over time to cause large disparities (Valian, 1999). Small gender differences in base salary at the start of one's career, for instance, can compound and multiply over time to produce large salary differences at later career stages. Similarly, minor instances of social exclusion and limited access to influential networks can accumulate to hinder women's advancement into positions of power and leadership (Clark & Corcoran, 1986; Hart, 2016; O'Meara et al., 2017). Using a computer simulation, Martell, Lane, and Emrich (1996) showed that when gender bias accounted for merely 1% of performance rating variance, women's promotion rates took a severe hit: only 35% of top positions went to women by the end of the simulation, compared to 50% when no bias was present. Subtle bias may seem minor and innocuous, but the cumulative effects can even be as detrimental as blatant discrimination.

Given the covert and unconscious nature of subtle bias, single incidents often go undetected and fail to arouse action to challenge their occurrence (Barreto & Ellemers, 2005). Scholars (e.g., Bird, 2011; Mitchneck, Smith, & Latimer, 2016; Moss-Racusin et al., 2014) have thus called for the development of interventions to educate individuals about the subtle nature of bias so they are better able to detect, report, and avoid bias. However, educating about bias can be challenging.

**Challenges to Teaching about Subtle Bias**

Teaching about gender bias poses many challenges that make the method of simply providing information insufficient. Being presented with evidence of injustice and bias can be

threatening and elicit negative reactions (Lerner & Miller, 1978; Tajfel & Turner, 1986). For instance, learning that gender bias is pervasive can threaten one's view of the world as fair and just (Major, Kaiser, O'Brien, & McCoy, 2007; Schmitt, Branscombe, & Postmes, 2003), provoking defensiveness and rejection of the information (Cundiff et al., 2014; Zawadzki, Danube, & Shields, 2012; Zawadzki et al., 2014). Furthermore, information that one may harbor unconscious biases can threaten one's egalitarian identity and lead one to reject negative feedback as a valid and accurate indicator of bias (Howell, Gaither, & Ratliff, 2015). Learning that gender bias is pervasive can also make individuals feel helpless to intervene and undermine their self-efficacy, that is, confidence in their ability to effect change (Cundiff et al., 2014; Zawadzki et al., 2012, 2014).

Given these challenges, it is no surprise that sexism interventions that rely on simply providing information to participants show mixed results for decreasing sexist attitudes and motivating behavior to address bias (Cundiff et al., 2014; Zawadzki et al., 2014). Yet, diversity training programs often rely on passive learning techniques to transmit information to recipients (Dobbin & Kalev, 2013; Moss-Racusin et al., 2014), ignoring the considerable literature that demonstrates the superiority of active and experiential learning strategies (i.e., methods that engage participants in discovery, reflection, and discussion) for improved learning outcomes (Prince, 2004). Effective interventions must present information in a way that minimizes defensiveness, promotes message acceptance, and instills a sense of self-efficacy that one has the ability and knowledge to combat sexism.

**Active and Experiential Learning as a Potential Solution**

Interventions must be able to increase awareness and recognition of subtle gender bias without eliciting defensiveness or belief that nothing can be done to remedy bias. One potential

strategy is to incorporate active and experiential learning methods (e.g., Kolb, 1984) that allow participants to discover knowledge about subtle gender bias on their own, rather than being passive recipients of information. WAGES (Shields et al., 2011) is one example of a bias-reduction intervention that employs active and experiential learning.

Drawing from experiential learning theory (Kolb & Kolb, 2005), WAGES engages participants in a game-like simulation that demonstrates how seemingly trivial gender biases can accumulate over time to negatively impact women in the workplace. Although bias is often thought of as obvious, much gender bias is unconscious meaning that most individuals—including faculty—lack experience with identifying when gender bias is occurring. WAGES provides participants with a hands-on experience where they can actively construct knowledge incrementally and through dialogue, which makes information about subtle gender bias less threatening and sustains participants' self-efficacy to address bias (Cundiff et al., 2014; Zawadzki et al., 2012, 2014). Experimental studies demonstrate that, relative to control conditions, playing WAGES increases knowledge about sexism (Shields et al., 2011; Zawadzki et al., 2012), reduces sexist attitudes (Zawadzki et al., 2014), and increases behavioral intentions to reduce sexism (Cundiff et al., 2014). More generally, experiential learning has been used effectively with adult populations, including skilled professionals (Fominykh, Leong, & Cartwright, 2017).

Prior work on WAGES, however, has focused on attitudinal outcomes. It is unknown whether WAGES motivates and equips people to *apply* their new knowledge to influence actual behavior and detection of bias when it occurs. Addressing this possible attitude-behavior gap is critical because attitudes do not always align with behavior (Ajzen, 2012). Several factors can prevent individuals from addressing bias despite desire and intention to do so, including feeling

that one's actions will be futile; that is, when they lack self-efficacy to effect change (Stewart, Latu, Branscombe, & Denney, 2010). The present study examines the impact of WAGES on actual behavior – the detection of subtle gender bias and engagement in bias-confronting behaviors. Because WAGES allows self-efficacy to remain intact (Cundiff et al., 2014; Zawadzki et al., 2012, 2014), we expect that WAGES will influence actual behavior to address bias.

**Study Overview and Hypotheses**

We employed a two-part study design to examine the effectiveness of WAGES for increasing the detection and reporting of subtle gender bias, also referred to as subtle sexism. Part one comprised the *intervention phase* in which participants either completed WAGES or one of two control conditions. Part two comprised the *application phase* that participants completed at least one week later as an ostensibly unrelated study. During the application phase, participants evaluated reviews of the promotion and tenure (P&T) materials of a woman candidate. Participants were randomly assigned to evaluate reviews that were either blatantly sexist, subtly sexist, or nonsexist. The two-part study design allowed us to examine whether participants would spontaneously apply the knowledge they had learned in WAGES to identify and report subtle sexism (but not erroneously report non-sexism) in a novel context.

We hypothesized that WAGES would influence the detection of bias (H1) and reporting of bias (H2) in the subtle sexism condition, but not in the blatant or nonsexist conditions. Subtle sexism can be difficult to detect, and so we predicted that those who had completed WAGES would be better able to detect and more likely to report subtle bias compared to control participants. Blatant sexism, by contrast, is relatively easy to detect (Swim et al., 2005), and so we did not expect differences in bias detection and reporting between WAGES participants and

control participants – we expected all participants to be equally adept at detecting and reporting blatant bias, regardless of whether they completed WAGES or not. Likewise, in the nonsexist condition, we expected detection and reporting of bias to be relatively low and similar across WAGES and control participants, thus demonstrating that playing WAGES does not over-sensitize participants to potential bias.

In addition, we also made predictions about the mechanism by which WAGES influences bias-reporting behavior. We predicted that the reason WAGES participants would be more likely to report bias in the subtle sexism condition is precisely because they would detect more bias compared to control participants (H3).

## Method

### Participants

Undergraduate students were recruited from the Psychology Department subject pool at a large mid-Atlantic university to participate in the intervention phase. Students who completed the intervention phase ($N$ = 363) were recruited approximately one week later for an ostensibly unrelated study, the application phase, of which 192 participated. Attrition rates did not vary across intervention conditions, $X^2$ (2, $N$ = 363) = 3.04, $p$ = .219. All participants received course credit. We aimed to recruit as many participants as possible in a single semester; data collection closed once the semester ended. Participants were excluded from analyses if, at the application phase, they did not correctly identify the gender of both the candidate and reviewer ($n$ = 7), expressed suspicion about the study ($n$ = 5), or did not complete the questions that probed for suspicion ($n$ = 3). Analyses included 177 participants: 139 self-identified as women, 36 as men, and 2 did not indicate their gender ($M_{age}$ = 18.94, $SD$ = 2.26). Most identified as white or Caucasian (85.9%), followed by black or African American (4.5%), and bi- or multi-racial

(3.4%). Although we did not determine sample size a priori, a sensitivity analysis using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007), with power set at .80 and alpha set at .05, indicated that the present sample (N = 177) was large enough to detect medium-sized effects (*f* = .26 or larger), which have been found in prior WAGES research (e.g., Shields et al., 2011; Zawadzki et al., 2012, 2014).

**Procedure and Materials**

**Intervention phase**. Participants arrived to the lab in groups of four to 12 where they were greeted by a woman experimenter. They were randomly assigned to either the WAGES condition (*n* = 55), an information-only condition (*n* = 60), or a group activity condition (*n* = 62). Each condition is described below. Immediately following the intervention, participants responded to a series of measures extraneous to the present study that are reported elsewhere (see Cundiff et al., 2014).

***WAGES condition.*** WAGES engages participants in a game-like simulation of subtle sexism in the academic workplace (Shields et al., 2011). Participants are divided into two teams, a white team and a green team. The object of the game is to be the first person to reach the end of the academic career ladder, progressing from Assistant Professor to Distinguished Professor. Players move through the game by drawing cards that award credit chips as well as how many spaces to move ahead. Players must both advance on the game board and earn enough credit chips to be promoted. The same standards apply to all players, but the accumulation of moves and chips depends on the game cards, which differ between the two teams. Each card depicts a realistic scenario drawn from empirical evidence of women's and men's experiences in the academic workplace. The white and green teams encounter the same scenarios (e.g., "A senior faculty member congratulates you on completing a big grant proposal on time"), but the

outcomes of the scenarios vary slightly between the two teams in a way that gives small cumulative advantages to the white team (e.g., the faculty member attributes the successful grant submission to skill for white team members, worth two chips and one move, but luck for green team members, worth one chip and one move). The differential outcomes experienced by the white and green teams are based on empirical research evidence of subtle sexism. Although the white and green teams encounter the same scenarios, the cards are shuffled within each pile so that the teams do not encounter the same scenario at the exact same time. This way, the inequitable nature of outcomes is subtle and not readily apparent at the outset because comparative information is lacking, similar to how subtle bias operates in the real world. But as the game progresses and the comparative scenarios are revealed, participants discover that slight advantages are being given to the white team. These slight advantages seem trivial at first— worth only one credit chip at a time— but accumulate over the course of the game to create large disparities between the two teams. The game portion ends once someone reaches Distinguished Professor, and then a facilitator leads a guided discussion that makes WAGES' learning objectives clear. In line with the process of experiential learning (Kolb, 1984), the facilitator guides participants to reflect on their experience and form abstract concepts about the meaning of the experience. Over the course of the game and through facilitated discussion and reflection, participants discover connections between the game and real-world experiences of subtle sexism. Discussion concludes with thoughts from participants about how to apply the information to recognize and address subtle sexism.

***Information-only condition.*** Participants in the information-only condition received all of the information contained in WAGES, but in a passive format. Specifically, participants read handouts about the subtle sexism examples illustrated in the WAGES game cards and a transcript

of the post-game discussion. The WAGES and information-only conditions were thus equally informative, but varied in presentation format, providing a test of the active learning component of WAGES.

***Group-activity condition.*** We also included a second control condition, a group activity in which participants played a modified version of Chutes and Ladders® and engaged in a discussion about workgroup dynamics with no explicit focus on subtle gender bias. Modifications to the game were minor with participants ostensibly playing on a green or white team, but otherwise playing as individuals according to the standard rules. The discussion focused on how to improve cooperation within groups, such as having common goals, equal status, and potential for friendly interactions. This control condition was designed to be similar to WAGES in terms of engagement and the active learning format, but without information about subtle sexism.

**Application phase**. Approximately one week after the intervention phase, participants came to a different lab in groups of up to seven where they were greeted by a different woman experimenter and were assigned to an individual computer station. Participants were told that the University Promotion and Tenure (P&T) Committee is interested in involving students in faculty promotion and tenure decisions. Thus, they would read an evaluation of a professor and would give their impressions of the candidate, the reviewer, and the review process. Participants were randomly assigned to read one of three evaluations, which provided the context for the blatant, subtle, or nonsexist manipulations. The left column of Table 1 shows the number of participants in each condition.

After receiving instructions, participants read a “Candidate Evaluation Form” for an ostensibly real faculty member, Jennifer Watkins. The form contained two types of information:

numeric performance ratings provided by the P&T Committee on 5-point scales (1 = *seriously below expectations*, 2 = *below expectations*, 3 = *meets expectations*, 4 = *exceeds expectations*, 5 = *outstanding*), and a written review submitted by a committee member, Greg Harris. We used the numeric and written review format to reflect real-world performance evaluations (Biernat, Tocci, & Williams, 2012). The numeric ratings and written review evaluated the candidate across three domains: research, teaching, and interpersonal and leadership skills. Participants read that both numeric ratings and written reviews are important for P&T decisions, and were given instructions for interpretation: "Candidates who receive tenure tend to have ratings of 3 (*meets expectations*) or higher in all domains and tend to have generally positive comments provided by the reviewer."

Numeric ratings were held constant across sexism type and indicated that the candidate met expectations for tenure. However, the reviewer comments varied and were the basis for the sexism type conditions. Specifically, the reviewer made similar observations across conditions, such as noting that "the candidate missed a few classes last year," but then provided different interpretations. In the nonsexist condition, the reviewer did not interpret his observations in light of the candidate's gender, commenting for example, "she always found a colleague to cover for her." In the subtle sexism condition, the reviewer's interpretations were influenced by gender stereotypes, stating for example that the candidate's absences were "presumably due to family obligations. This raises some questions as to her dedication to her academic commitments." In the blatant sexism condition, the reviewer made explicit gender-based comparisons, such as comparing her with "male faculty [who] seldom prioritize family commitments over academic commitments." The P&T review materials were pretested such that the blatant sexist condition

was perceived as most biased by a separate group of undergraduates ($N$ = 27), followed by the subtle sexist and nonsexist conditions ($p$s < .03).

Next, participants completed the dependent measures, followed by an additional measure irrelevant to the present investigation and reported elsewhere (see Cundiff et al., 2014). Finally, participants were probed for suspicion and verbally debriefed.

**Measures**

Participants indicated their perceptions of the reviewer and the review process, which were embedded in a longer list of filler items that were included to give legitimacy to the cover story. For instance, participants rated the candidate (e.g., "Overall, I would recommend that this candidate be promoted/given tenure"), the review process (e.g., "I would have liked more information about this candidate in order to provide my opinion"), and indicated their interest in serving as a student representative on the P&T committee. These filler items were not central to hypotheses and therefore were not analyzed.

**Perceived negativity of review**. Participants rated the negativity of the review to ensure that the blatant sexist review was indeed perceived as more negative than the subtle and nonsexist reviews, which were also expected to differ. Using a 7-point scale (1 = *strongly disagree*, 7 = *strongly agree*), participants indicated their agreement with two items: "The reviewer made an overall negative assessment of the candidate" and "The reviewer feels positively towards the candidate" (reverse scored). We averaged across these two items; higher scores indicated perceptions of a more negative evaluation, $r(175) = .81$, $p < .001$.

**Detected bias.** Detected bias was assessed with three self-report measures that varied in terms of how directly they measured bias detection. Two measures were indirect – they assessed

anger reactions and perceived unfairness of the review. The third measure was direct – it explicitly asked participants to rate their perceptions of bias.

***Anger reactions.*** Because perceiving bias often elicits anger (Ellemers & Barreto, 2009), we assessed the extent to which the reviews made participants feel angry. Participants indicated on 7-point scales (1 = *not at all*, 7 = *very much*) the extent to which the review made them feel angry, upset, irritated, hostile, furious, and frustrated. These items were embedded within filler items (e.g., bored). Scores for the anger items were averaged; higher scores indicated stronger anger reactions ($\alpha = .95$).

***Perceived unfairness of the review.*** Participants rated their agreement with the following items using a 7-point scale (1 = *strongly disagree*, 7 = *strongly agree*): (1) "The reviewer seemed more critical of the candidate than was warranted", (2) "The reviewer's comments about the candidate were unfair at times", (3) "This reviewer provided a quality assessment of the candidate" (reverse-scored), and (4) "The reviewer made a fair evaluation of the candidate" (reverse-scored). These items measured concerns with fairness without explicitly drawing attention to gender bias. Responses were averaged; higher scores indicated stronger perceived unfairness ($\alpha = .94$).

***Explicit ratings of bias.*** Using 5-point scales (1 = *not at all [discriminatory, sexist, etc.]*, 5 = *very [discriminatory, sexist, etc.]*; adapted from O'Brien, Kinias, & Major, 2008), participants indicated the extent to which they thought: (1) the reviewer was discriminatory towards the candidate; (2) the reviewer was sexist in his evaluation of the candidate; (3) the reviewer's evaluation was influenced by the candidate's gender; and (4) the review accurately reflected the candidate's qualifications (reverse scored). Responses were averaged; higher scores indicated greater detected bias ($\alpha = .94$).

**Reported bias.** Participants were informed that the P&T Committee strives for fairness and equity in the review process, and to that end they were invited to anonymously report any concerns by responding yes or no to the following statement: "I have concerns that this review was unfair and would like to officially report my concerns to the Committee." Participants who indicated "yes" were invited to provide a description of their concerns to the Committee. Participants were led to believe that any reported concerns would be forwarded anonymously to the University P&T Committee.

## Results

Table 1 presents the means and standard deviations for all dependent measures.

### Perceived Negativity of Review

We conducted a one-way ANOVA to test that participants perceived the blatant sexist review to be the most negative, followed by the subtle and nonsexist reviews. Confirming the effectiveness of the sexism manipulation, there was a main effect of sexism type, $F(2, 174) = 105.40$, $p < .001$, $\eta_p^2 = .55$. Pairwise comparisons using Least Significant Difference revealed that participants perceived the review to be more negative in the blatant sexism condition than in the subtle ($p < .001$) and nonsexist ($p < .001$) conditions, which also differed ($p < .001$) (see Table 1, left column).

### H1: Effects of WAGES on Detected Bias

We first submitted each measure of detected bias to a 3 (intervention condition: WAGES, information-only, group activity) × 3 (sexism type: blatant, subtle, nonsexist) between-participant ANOVA. Recall that we predicted differences between WAGES and control participants in the subtle sexism condition, but not in the blatant or nonsexist conditions. Thus,

we examined planned contrasts to test for effects of intervention condition in each of the sexism conditions. Means and standard deviations are reported in the middle columns of Table 1.

***Anger reactions.*** The 3 x 3 ANOVA revealed a main effect of sexism type, $F(2, 168) = 38.05$, $p < .001$, $\eta_p^2 = .31$. The blatant sexist review made participants angrier compared with the subtle ($p < .001$) and nonsexist reviews ($p < .001$), which also differed ($p = .009$).

The predicted Intervention x Sexism Type interaction did not reach standard levels of significance, $F(4, 168) = 1.90$, $p = .113$, $\eta_p^2 = .04$. Planned comparisons revealed that, as predicted, there were no differences between intervention conditions in the blatant and nonsexist conditions, $F$s < 1. In the subtle sexism condition, however, significant differences emerged between the intervention conditions, $F(2, 168) = 3.56$, $p = .031$, $\eta_p^2 = .04$ (see Figure 1). Participants in the WAGES condition reacted to the subtle sexist review with more anger than participants in the group-activity condition ($p = .011$), and marginally more anger than participants in the information-only condition ($p = .073$). Information-only and group-activity participants did not differ in their anger reactions to the subtle sexist review, $p = .434$.

***Perceived unfairness of the review.*** The 3 x 3 ANOVA revealed a main effect of sexism type, $F(2, 168) = 92.72$, $p < .001$, $\eta_p^2 = .53$. Participants in the blatant sexism condition perceived more unfairness than those in the subtle ($p < .001$) and nonsexist ($p < .001$) conditions, which also differed ($p < .001$). Contrary to predictions, the Intervention x Sexism Type interaction was not significant, $F(4, 168) = 0.95$, $p = .440$, $\eta_p^2 = .02$. Planned contrasts did not reveal any significant differences between intervention conditions in any of the sexism conditions (see Figure 2).

***Explicit ratings of bias.*** The 3 × 3 ANOVA revealed a main effect of sexism type, $F(2, 168) = 118.75$, $p < .001$, $\eta_p^2 = .59$. Participants in the blatant sexism condition perceived more

bias than those in the subtle ($p < .001$) and nonsexist ($p < .001$) conditions, which also differed ($p < .001$). This main effect was qualified by the predicted Intervention × Sexism Type interaction, $F(4, 168) = 2.46$, $p = .047$, $\eta_p^2 = .06$ (see Figure 3). As predicted, there were no differences between intervention conditions in the blatant and nonsexist conditions, $F$s < 1. However, in the subtle sexism condition, there were significant intervention differences, $F(2, 168) = 6.56$, $p = .002$, $\eta_p^2 = .07$. Participants in the WAGES condition detected more bias than those in either the information-only ($p = .016$) or group activity ($p = .001$) conditions, which did not differ ($p = .277$).

**H2: Effects of WAGES on Reported Bias**

We next examined whether participants reported a concern (coded no = 0, yes = 1) to the P&T Committee. Figure 4 displays the proportion of participants in each condition who reported bias (see also Table 1, right column). Due to the dichotomous nature of the outcome variable, we used Generalized Linear Modeling and specified a binomial logistic model with a logit link. We used a 3 (intervention) × 3 (sexism type) between-participant design. The omnibus test was significant, $\chi^2(8) = 54.40$, $p < .001$. Examination of the parameter estimates revealed that the interaction term was not significant, $p > .45$; thus, we trimmed the model by removing the interaction term. The omnibus test for the trimmed model was also significant, $\chi^2(4) = 48.36$, $p < .001$. As shown in Table 2, there were significant main effects for intervention and sexism type. Pairwise comparisons for intervention (see top half of Table 2) revealed that WAGES participants were more likely to report concerns compared to group-activity and information-only participants, which did not differ. Pairwise comparisons for sexism type (see bottom half of Table 2) revealed that participants in the blatant sexism condition were more likely to report

concerns compared to participants in the subtle sexism and nonsexist conditions, which also differed.

We next conducted the planned tests for intervention effects in the subtle sexism condition only by selecting participants in the subtle sexism condition and using Generalized Linear Modeling to test for intervention condition differences. The omnibus test was significant, $\chi^2(2) = 7.76$, $p = .021$. Consistent with predictions, there was a significant intervention effect (see top row of Table 3). As predicted, participants in the WAGES condition were more likely to report concerns about the subtle sexist review than were information-only and group activity participants, who did not differ (see Table 3 and Figure 4). We also checked for intervention effects in the blatant sexism and nonsexist conditions; no significant intervention differences emerged ($p$s > .410; see Figure 4). Thus, as predicted, only the subtle sexism condition revealed differences between the intervention conditions.

**H3: Detected Bias as a Mediator**

We conducted a mediation analysis using PROCESS (model 4; Hayes, 2013) to test the hypothesis that within the subtle sexism condition, participants who completed WAGES would detect more subtle bias (explicit measure of detected bias) and subsequently be more likely to report concerns than participants in the two control conditions. The control conditions were combined because these participants did not differ in detected bias or reported concerns. We followed the bootstrapping procedure described by Preacher and Hayes (2008) and estimated confidence intervals using 5,000 bootstrap re-samples. Evidence of mediation would be established if intervention condition (coded as 0 = controls, 1 = WAGES) predicted detected bias, detected bias predicted reporting concerns (controlling for intervention condition), and the confidence interval constructed for this indirect effect did not include zero.

The overall model was significant: $F(1, 53) = 8.41$, $p = .005$, $R^2 = .14$. In line with the results reported above, WAGES participants detected more subtle bias than control participants: $b = .81$, $t = 2.90$, $p = .005$. Detected bias predicted reported concerns, controlling for intervention condition: $b = 1.10$, $Z = 2.56$, $p = .010$. The bias-corrected and accelerated confidence interval for the indirect effect did not include zero, 95% CI [.17, 2.56], indicating that detected bias mediated the relationship between intervention condition (WAGES vs. controls) and reported concerns. Further, results indicated full mediation, as the effect of intervention condition on reported concerns was no longer significant when detected bias was included the model (total effect: $b = 1.85$, $Z = 2.65$, $p = .008$; direct effect: $b = 1.33$, $Z = 1.74$, $p = .082$).

**Exploratory Analyses**

We also considered the possibility that participants' preexisting sexist attitudes could explain our findings if those attitudes systematically varied with experimental condition. To test this possibility, we accessed participants' scores on the Perceived Harm of Everyday Sexism scale (PHES; Cundiff et al., 2014), which was administered prior to this study in a separate session as a part of an unrelated survey (see Cundiff et al., 2014). We submitted PHES scores to a 3 (intervention condition) x 3 (sexism type) ANOVA. No effects were significant, $F$'s $\leq 1.00$, $p$s $\geq .408$, confirming that random assignment to conditions was successful. We also reran all analyses controlling for PHES scores. PHES scores were not a significant covariate in any of the analyses, and the pattern of results did not change for any of the variables.

## Discussion

Consistent with hypotheses, participants who completed WAGES (vs. controls) were better able to detect and report subtle gender bias in the P&T materials. In addition, mediation analyses revealed that the ability of WAGES (vs. controls) to enhance the detection of subtle bias

was a key mechanism through which WAGES exerted its effect on reporting subtle bias. That is, WAGES participants were more likely to report subtle bias compared to control participants precisely because they were better able to detect subtle bias.

We assessed bias detection using three measures that varied in directness. The expected pattern of results emerged for two of the three measures: when presented with subtle sexism, WAGES participants felt angrier and, when probed directly, reported perceiving more bias than did control participants. When indirectly probed about gender bias, however, WAGES and control participants reported similar levels of perceived unfairness for the subtle sexist review. It is unclear why our measure of perceived unfairness did not yield the predicted differences between WAGES and other two conditions. One possibility is that perceived unfairness was less precise in measuring bias detection, potentially conveying that the reviewer's standards were overly critical, as well as standards being gender biased. Another possibility is that people are more willing to label behavior as unfair than as sexist, and that educating about subtle sexism helps remove this barrier. Indeed, only those in the WAGES condition felt angrier and were more willing to directly identify the review as sexist, which are both important precursors to taking action to address bias (Ellemers & Barreto, 2009).

Notably, participants in all intervention conditions were able to detect and report blatant sexism. This finding is encouraging and suggests that people can easily identify and report overt forms of sexism without the need for intervention. Subtle sexism, by contrast, is more difficult to detect. Our results suggest that without appropriate intervention, individuals are unlikely to detect subtle forms of gender bias, and as a result, are unlikely to take action to address bias. These results support other work (e.g., Bird, 2011; Mitchneck et al., 2016; O'Meara et al., 2017) in highlighting the need for interventions that educate about subtle forms of gender bias. Further,

participants in all three intervention conditions detected equivalent and minimal bias in the nonsexist P&T materials, indicating that WAGES did not over-sensitize participants to bias.

Interestingly, receiving information about gender bias in a passive learning format (information-only condition) produced the same results as receiving no information about gender bias (group-activity condition), suggesting that simply providing information about gender bias is not enough to facilitate subsequent bias detection and reporting. Instead, presenting information in a way that actively engages participants to discover knowledge on their own (WAGES condition), rather than passively, produced superior learning outcomes and better equipped participants to detect and report subtle bias. These findings complement prior work on WAGES (Cundiff et al., 2014; Shields et al., 2011; Zawadzki et al., 2012, 2014) and demonstrates the superiority of active and experiential learning strategies over passive learning strategies for teaching about gender bias. Specifically, compared to passive teaching methods (e.g., information-only condition in the present study), prior work shows that active and experiential learning (as in WAGES) produces superior outcomes because it minimizes resistance to learning about gender bias, enabling learners to be more receptive to the information. Learning about bias can threaten core psychological needs, such as needs to control one's outcomes, belong and affiliate with others, and maintain positive views of the self (Lerner & Miller, 1978; Tajfel & Turner, 1986). Information about gender bias must thus be presented in a way that is non-threatening; our results suggest that employing active and experiential learning when teaching about bias is an effective way to minimize the threats associated with learning about bias.

More broadly, our research answers the call for the development and rigorous testing of interventions aimed at improving gender equity through training about bias. Women faculty

continue to face barriers to their career advancement. Whereas overt forms of bias have significantly decreased, subtle forms persist, stalling women faculty's career progress. One reason that subtle bias persists is because it is difficult to detect and thus fails to arouse action to mitigate it. Our study contributes to the literature on gender inequity in academia by demonstrating that interventions for educating about subtle gender bias may be most effective when they incorporate experiential learning strategies.

**Implications and Applications**

WAGES is not intended to be a magic bullet. The causes for gender disparities in academe are complex and will require multi-pronged approaches to remedy. WAGES is but one tool among many that are needed to fully achieve gender equity. Other measures are needed, including cultural changes regarding ideal worker norms (Sallee, Ward, & Wolf-Wendel, 2016), changes to family leave policies that remove stigma associated with taking leave (Ward & Wolf-Wendel, 2005), and collecting and sharing data on relevant metrics such as salary and workload inequity (O'Meara et al., 2017). In order to garner support for change measures, however, individuals must first perceive and acknowledge the barriers that exist. Thus interventions that educate about bias, such as WAGES, may be an important first step toward motivating action.

Notably, bias-reduction training may garner limited participation if it is voluntary. To ensure high levels of participation, administrators may consider integrating bias-reduction interventions into existing training programs. For example, WAGES could be implemented within required graduate training as part of graduate student orientation or during a professional development course. WAGES could also be implemented during orientation sessions for new faculty and/or required for hiring and P&T committees.

Extending our results further, incorporating active and experiential learning strategies into bias-reduction interventions more broadly may be a fruitful endeavor. Similar intervention procedures could be employed to educate individuals regarding other forms of subtle bias, such as bias based on race-ethnicity, sexual orientation, class, or ability status. Active and experiential learning may prove to be a valuable tool for those seeking to educate about bias across a variety of domains.

**Limitations and Future Directions**

WAGES was effective in a novel, real-world situation that was ostensibly unrelated to the intervention itself and occurred at least a week later. Yet, one limitation to the generalizability of our findings is our sample of undergraduate students. On one hand, the fact that WAGES was effective in empowering students—a population that likely has limited knowledge and perhaps minimal interest in faculty career advancement—speaks to the power of the intervention. On the other hand, it is plausible that the experiential learning format of WAGES is less effective for faculty and administrators—its intended audience—than undergraduate students, who may be more open to attitude change than faculty. However, prior work indicates that experiential learning is effective not only for undergraduates (Prince, 2004), but also for skilled adult professionals (Fominykh et al., 2017). Additionally, recent tests of WAGES with faculty and administrator samples show similar attitude changes as those reported with undergraduate samples, with similar indications of receptiveness and engagement (Shields, McCormick, Dicicco, & Zawadzki, in press). Nonetheless, future work should also test for behavioral changes with faculty and administrator samples to confirm generalizability of results.

Another limitation is that our sample consisted primarily of women. This limited our ability to test whether gender moderated our findings, particularly to test whether WAGES

would produce similar effects for a larger sample of men. Prior work, however, indicates that WAGES tends to be effective for both women and men, although at times WAGES produces stronger effects for women (Zawadzki et al., 2012, 2014). Future work is needed to further test the generalizability of findings to men, especially given that men outnumber women in positions that have power to make P&T decisions (Catalyst, 2015; Curtis, 2011). More generally, further research is needed to determine if, when, and why people of different genders might respond differently to sexism interventions.

Our experimental design provided participants with ample opportunity to report bias. In real-world settings, however, opportunities for reporting bias may be less clear. Policies and procedures for reporting bias should be made obvious and low-cost for reporters in terms of time, effort, and risk of repercussion in order to maximize bias reporting. For instance, universities could administer anonymous surveys at the conclusion of each P&T and/or hiring meeting that would give faculty ample opportunity to report bias anonymously.

The current research focused on WAGES-Academic, which is but one version of WAGES. Other versions are currently being developed, including a classroom version focused on undergraduate experiences. Future work could adapt WAGES to other work contexts, such as careers in corporate business or engineering.

Future work could also adapt WAGES to educate about other forms of bias. WAGES focused on gender, but could be further developed to focus on other marginalized identities and their intersections. Likewise, in discussing gender, WAGES primarily used a binary framework that does not account for the experiences of those who identify outside this binary. Many of the items relating to token status, exclusion, and double standards (among others) are relevant to

other marginalized groups. Yet, future work with WAGES and other interventions would benefit from considering multiple identities to broaden the impact of these efforts.

**Conclusion**

Without proper intervention, subtle forms of gender bias are difficult to detect. However, information about subtle bias must be framed appropriately in order to produce desired effects. Merely presenting information about gender bias is not enough. Rather, our results demonstrate the superiority of using active and experiential learning over traditional passive learning formats. By providing participants with a relevant experience on which to reflect and discover knowledge on their own, participants were better able to detect and more likely to report subtle bias in an unrelated task one week later. Overall, our results suggest that even a brief, simple intervention can enhance detection of subtle bias. WAGES is one example of an effective low-cost intervention; other similar interventions also show promise (e.g., Carnes et al., 2012; Smith, Handley, Zale, Rushing, & Potvin, 2015). The key point is that interventions designed to educate about bias are likely to be most effective when they actively engage participants to discover knowledge on their own, reflect on the experience, and discuss their experience with others. Simply providing information to passive recipients is not enough to effect change.

Acknowledging that bias exists is an important first step toward addressing it. But without proper intervention, acknowledgment and detection of bias is unlikely. Institutional change starts with changing the attitudes and behavior of individuals within the institution (Lindquist, 1974; Simpson & Flynn, 2007). By increasing awareness of bias in higher education institutions in a non-threatening way, WAGES and other similar interventions that employ active and experiential learning can be used as a starting point to effectively change attitudes and

behaviors so that individuals may be more likely to instigate and adopt the institutional and structural changes needed to fully address gender inequity (Sevo & Chubin, 2008).

References

Ajzen, I. (2012). The theory of planned behavior. In P. M. Van Lange, A. W. Kruglanski, E. T. Higgins, P. M. Van Lange, A. W. Kruglanski, E. T. Higgins (Eds.), *Handbook of theories of social psychology*, (Vol. 1, pp. 438-459). Thousand Oaks, CA: Sage Publications Ltd. doi:10.4135/9781446249215.n22

Barreto, M., & Ellemers, N. (2005). The perils of political correctness: Men's and women's responses to old-fashioned and modern sexist views. *Social Psychology Quarterly, 68*, 75-88. doi: 10.1177/019027250506800106

Biernat, M., Tocci, M. J., & Williams, J. C. (2012). The language of performance evaluations: Gender-based shifts in content and consistency of judgment. *Social Psychological and Personality Science, 3*, 186-192. doi: 10.1177/1948550611415693

Bird, S. R. (2011). Unsettling universities' incongruous, gendered bureaucratic structures: A case-study approach. *Gender, Work and Organization, 18*, 202–230. doi:10.1111/j.1468-0432.2009.00510.x

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*(2), 63-77. doi: 10.1037/a0028128

Catalyst. (2015). *Quick take: Women in academia*. New York: Catalyst. Available at http://www.catalyst.org/knowledge/women-academia#us

Clark, S. M., & Corcoran, M. (1986). Perspectives on the professional socialization of women faculty: A case of accumulative disadvantage?. *The Journal of Higher Education, 57*, 20-43.

Cooper, J., Eddy, P., Hart, J., Lester, J., Lukas, S., Eudey, B., & ... Madden, M. (2007). Improving gender equity in postsecondary education. In S. S. Klein, B. Richardson, D. A. Grayson, L. H. Fox, C. Kramarae, D. S. Pollard, ... C. A. Dwyer (Eds.), *Handbook for achieving gender equity through education*, 2nd ed (pp. 631-653). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Cress, C. M., & Hart, J. (2009). Playing soccer on the football field: The persistence of gender inequities for women faculty. *Equity & Excellence in Education*, *42*(4), 473-488.

Crosby, F., Clayton, S., Alksnis, O., & Hemker, K. (1986). Cognitive biases in the perception of discrimination: The importance of format. *Sex Roles*, *14*, 637-646.

Cundiff, J. L., Zawadzki, M. J., Danube, C. L., & Shields, S. A. (2014). Using experiential learning to increase the recognition of everyday sexism as harmful: The WAGES intervention. *Journal of Social Issues*, *70*(4), 703-721. doi: 10.1111/josi.12087

Curtis, J. W. (2011). *Persistent inequity: Gender and academic employment*. Prepared for "New Voices in Pay Equity", An Event for Equal Pay Day (Association of American College and Universities), April 11, 2011. Available at https://www.aaup.org/NR/rdonlyres/08E023AB-E6D8-4DBD-99A0-24E5EB73A760/0/persistent_inequity.pdf

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5-18. doi: 10.1037/0022-3514.56.1.5

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

Ellemers, N., & Barreto, M. (2009). Collective action in modern times: How modern expressions of prejudice prevent collective action. *Journal of Social Issues, 65*(4), 749-768.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi: 10.3758/BF03193146

Fominykh, M., Leong, P. & Cartwright, B. (2017). Role-playing and experiential learning in a professional counseling distance course. In J. Johnston (Ed.), *Proceedings of EdMedia 2017* (pp. 1078-1090). Washington, DC: Association for the Advancement of Computing in Education (AACE).

Hart, J. L. (2016). Dissecting a gendered organization: Implications for career trajectories for mid-career women in STEM. *Journal of Higher Education, 87*, 605–634. doi:10.1353/jhe.2016.0024

Hart, J. L., & Cress, C. M. (2008). Are women faculty just "worrywarts?" Accounting for gender differences in self-reported stress. *Journal of Human Behavior in the Social Environment, 17*(1-2), 175-193.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior, 32*, 113-135.

Heilman, M. E., & Haynes, M. C. (2008). Subjectivity in the appraisal process: A facilitator of gender bias in work settings. In E. Borgida & S. T. Fiske (Eds.), *Beyond common sense: Psychological science in the courtroom* (pp. 127-155). Malden: Blackwell Publishing.

Hirshfield, L. E., & Joseph, T. D. (2012). ‘We need a woman, we need a black woman’: Gender, race, and identity taxation in the academy. *Gender and Education, 24*, 213-227.

Howell, J. L., Gaither, S. E., & Ratliff, K. A. (2015). Caught in the middle: Defensive responses to IAT feedback among whites, blacks, and biracial black/whites. *Social Psychological and Personality Science, 6*(4), 373-381. doi: .1177/1948550614561127

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58,* 697–720. doi: 10.1037/0003-066X.58.9.697

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Kolb, A. Y., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning and Education, 4*, 193-212. doi:10.5465/AMLE.2005.17268566

Lerner, M. J., & Miller, D. T. (1978). Just world research and the attribution process: Looking back and ahead. *Psychological Bulletin, 85*, 1030–1051. doi:10.1037/0033-2909.85.5.1030

Lester, J. (2011). Regulating gender performances: Power and gender norms in faculty work. *NASPA Journal about Women in Higher Education, 4*, 142-169.

Lindquist, J. (1974). Political linkage: The academic innovation process. *The Journal of Higher Education, 45*, 323–343.

Livingston, R. W. (2013). Gender, race, and leadership: An examination of the challenges facing non-prototypical leaders. Paper presented at the *Gender & Work Research Symposium: Challenging Conventional Wisdom*, Boston, MA.

Major, B., Kaiser, C. R., O'Brien, L. T., & McCoy, S. K. (2007). Perceived discrimination as worldview threat or worldview confirmation: implications for self-esteem. *Journal of Personality and Social Psychology*, *92*(6), 1068-1086. doi: 10.1037/0022-3514.92.6.1068

Marschke, R., Laursen, S., Nielsen, J. M., & Rankin, P. (2007). Demographic inertia revisited: An immodest proposal to achieve equitable gender representation among faculty in higher education. *The Journal of Higher Education, 78*, 1-26.

Martell, R. F., Lane, D. M., & Emrich, C. (1996). Male-female differences: A computer simulation. *American Psychologist, 51*, 157-158.

Martínez Alemán, A. M. (2014). Managerialism as the "new" discursive masculinity in the university. *Feminist Formations, 26*, 107-134. doi:10.1353/ff.2014.0017

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: http://dx.doi.org/10.1037/apl0000022

Monroe, K., Ozyurt, S., Wrigley, T., & Alexander, A. (2008). Gender equality in academia: Bad news from the trenches, and some possible solutions. *Perspectives on Politics*, *6*(2), 215-233. doi: 10.1017/S1537592708080572

Mitchneck, B., Smith, J., & Latimer, M. (2016). A recipe for change: Creating a more inclusive academy. *Insights, 352*(6282), 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*, 16474-16479. doi: 10.1073/pnas.1211286109

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*(6171), 615-616.

Muhs, G. G. & Niemann, Y. F. & González, C. G. & Harris, A. P. (2012). *Presumed incompetent: The intersections of race and class for women in academia*. Logan, UT: Utah State University Press.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., Smith, C. T., Olson, K. R., Chugh, D., Greenwald, A. G., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*(1), 36-88. doi: 10.1080/10463280701489053

O'Brien, L. T., Kinias, Z., & Major, B. (2008). How status and stereotypes impact attributions to discrimination: The stereotype-asymmetry hypothesis. *Journal of Experimental Social Psychology*, *44*, 405-412. doi 10.1016/j.jesp.2006.12.003

O'Meara, K. Kuvaeva, A., Nyunt, G., Jackson, R. & Waugaman, C. (2017). Asked More Often: Gender Differences in Faculty Workload in Research Universities and the Work Interactions that Shape Them. *American Educational Research Journal*, 1–33. doi: 10.3102/0002831217716767

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879-891. doi: 10.3758/BRM.40.3.879

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223-231. doi: 10.1002/j.2168-9830.2004.tb00809.x

Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, *28*, 61-79.

Sallee, M., Ward, K., & Wolf-Wendel, L. (2016). Can anyone have it all? Gendered views on parenting and academic careers. *Innovative Higher Education, 41*(3), 187-202.

Schmitt, M. T., Branscombe, N. R., & Postmes, T. (2003). Women's emotional responses to the pervasiveness of gender discrimination. *European Journal of Social Psychology*, *33,* 297-312. doi: 10.1002/ejsp.147

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

Shields, S. A., McCormick, C., Diciccio, E., & Zawadzki, M. J. (in press). Demonstrating the cumulative effects of unconscious bias with WAGES-Academic (Workshop Activity for Gender Equity Simulation): Short and long-term impact on faculty and administrators. *Journal of Women and Minorities in Science and Engineering*.

Shields, S. A., Zawadzki, M. J., & Johnson, R. N. (2011). The impact of the Workshop Activity for Gender Equity Simulation in the Academy (WAGES-Academic) in demonstrating cumulative effects of gender bias. *Journal of Diversity in Higher Education*, *4*, 120-129. doi: 10.1037/a0022953

Simpson, D. D., & Flynn, P. M. (2007). Moving innovations into treatment: A stage-based approach to program change. *Journal of Substance Abuse Treatment*, *33,* 111–120.

Smith, J. L., Handley, I. M., Zale, A. V., Rushing, S., & Potvin, M. A. (2015). Now hiring! Empirically testing a three-step intervention to increase faculty gender diversity in STEM. *BioScience*, *65*(11), 1084-1087. doi: 10.1093/biosci/biv138

Steinpreis, R. E., Anders, K. A., & Ritzke, D. (1999). The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles*, *41*, 509-528.

Stewart, T. L., Latu, I. M., Branscombe, N. R., & Denney, H. T. (2010). Yes we can! Prejudice reduction through seeing (inequality) and believing (in social change). *Psychological Science, 21*, 1557-1562.

Swim, J. K., Mallett, R., Russo-Devosa, Y., & Stangor, C. (2005). Judgments of sexism: A comparison of the subtlety of sexism measures and sources of variability in judgments about sexism. *Psychology of Women Quarterly, 29*, 406-411. doi: 10.1111/j.1471-6402.2005.00240.x

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of Intergroup Relations* (pp. 7-24). Chicago: Nelson-Hall.

Turner, C. S. V., Gonzalez, J. C., & Wong, K. (2011). Faculty women of color: The critical nexus of race and gender. *Journal of Diversity in Higher Education, 4*(4), 199–211.

U.S. Department of Education, National Center for Education Statistics. (2015). *Digest of Education Statistics*, Table 315.10. Available from http://nces.ed.gov/programs/digest/d14/tables/dt14_315.10.asp

Valian, V. (1999). *Why so slow? The advancement of women*. Cambridge, MA: The MIT Press.

Ward, K., & Wolf-Wendel, L. E. (2005). Work and family perspectives from research university faculty. *New Directions for Higher Education, 2005*(130), 67-80.

Wennerås, C., & Wold, A. (1997). Nepotism and sexism in peer review. *Nature, 387*(6631), 341-343.

Williams, J. C. (2014). Hacking tech's diversity problem. *Harvard Business Review, 92*(10), 94-100.

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-z

Zawadzki, M. J., Shields, S. A., Danube, C. L., & Swim, J. K. (2014). Reducing the endorsement of sexism using experiential learning: The Workshop Activity for Gender Equity Simulation (WAGES). *Psychology of Women Quarterly, 38*, 75-92. doi: 10.1177/0361684313498573

Table 1

*Means (standard deviations) of dependent measures across sexism and intervention conditions.*

| | *N* | Perceived Negativity | Anger Reactions | Perceived Unfairness | Bias Ratings | Reported Bias |
|---|---|---|---|---|---|---|
| Blatant Sexist | | | | | | |
| WAGES | 15 | 5.17 (1.03) | 3.18 (1.75) | 5.40 (1.37) | 3.98 (0.89) | 0.67 (0.49) |
| Information Only | 24 | 5.02 (1.49) | 3.33 (1.78) | 5.38 (1.26) | 4.02 (0.87) | 0.50 (0.51) |
| Group Activity | 26 | 2.04 (1.52) | 3.51 (1.79) | 5.53 (1.23) | 4.07 (0.75) | 0.62 (0.50) |
| *Overall* | *65* | *5.06 (1.39)* | *3.37 (1.75)* | *5.44 (1.25)* | *4.03 (0.81)* | *0.58 (0.50)* |
| Subtle Sexist | | | | | | |
| WAGES | 20 | 4.08 (1.23) | 2.58 (1.39) | 4.06 (1.46) | 3.16 (1.01) | 0.45 (0.51) |
| Information Only | 18 | 3.94 (1.14) | 1.81 (1.12) | 3.74 (1.19) | 2.50 (0.92) | 0.11 (0.32) |
| Group Activity | 17 | 3.24 (1.08) | 1.47 (0.62) | 3.40 (1.63) | 2.19 (1.07) | 0.12 (0.33) |
| *Overall* | *55* | *3.77 (1.19)* | *1.98 (1.19)* | *3.75 (1.43)* | *2.65 (1.07)* | *0.24 (0.43)* |
| Nonsexist | | | | | | |
| WAGES | 20 | 1.55 (0.92) | 1.25 (0.51) | 2.05 (1.03) | 1.64 (0.72) | 0.10 (0.31) |
| Information Only | 18 | 2.06 (0.98) | 1.13 (0.23) | 2.42 (1.12) | 1.71 (0.74) | 0.11 (0.32) |
| Group Activity | 19 | 2.11 (0.94) | 1.54 (0.97) | 2.39 (0.98) | 1.64 (0.51) | 0.00 (0.00) |
| *Overall* | *57* | *1.89 (0.96)* | *1.31 (0.66)* | *2.28 (1.04)* | *1.66 (0.65)* | *0.07 (0.26)* |

Table 2

*Main effects and pairwise comparisons for reported bias.*

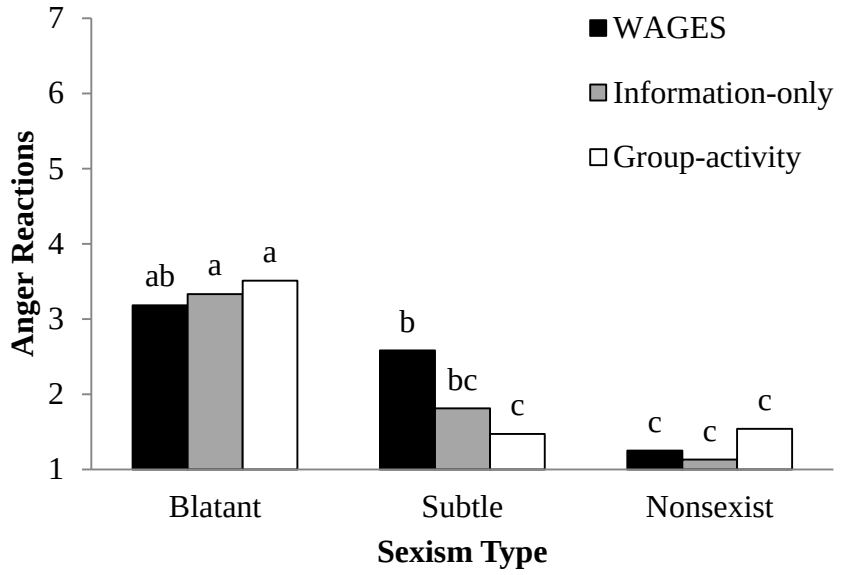
| | *B* | *SE* | *Wald χ2* | *p* |
|---|---|---|---|---|
| Intervention | | | 6.03 | .049 |
| WAGES vs. Info-only | 1.10 | 0.49 | 5.08 | .024 |
| WAGES vs. Group-activity | 0.99 | 0.48 | 4.22 | .040 |
| Info-only vs. Group-activity | 0.11 | 0.46 | 0.06 | .807 |
| Sexism Type | | | 33.34 | <.001 |
| Blatant vs. Subtle | 1.74 | 0.44 | 16.00 | <.001 |
| Blatant vs. Nonsexist | 3.19 | 0.61 | 27.65 | <.001 |
| Subtle vs. Nonsexist | 1.45 | 0.62 | 5.49 | .019 |

Table 3

*Planned comparison of reported bias between intervention conditions in the subtle sexism condition.*

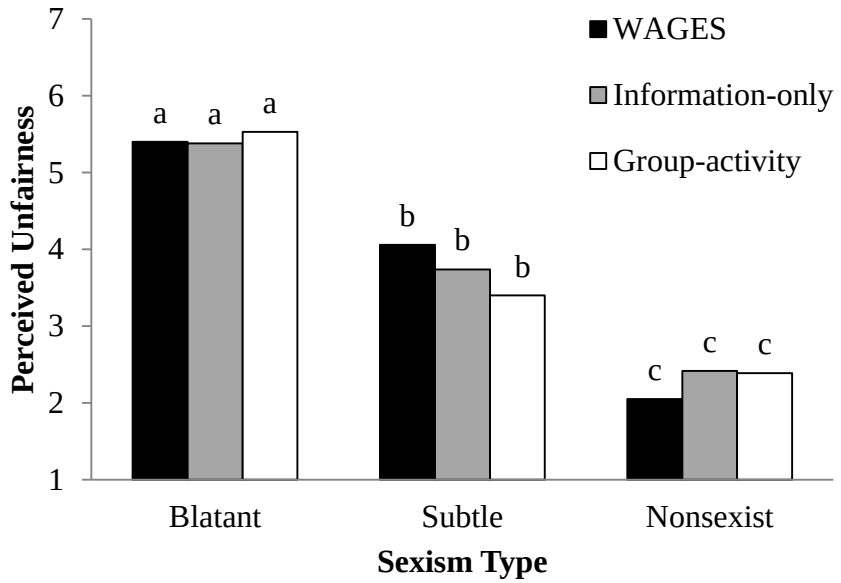
| | *B* | *SE* | *Wald* $\chi 2$ | *p* |
|---|---|---|---|---|
| Intervention (Subtle sexism only) | | | 7.05 | .030 |
| WAGES vs. Info-only | 1.88 | 0.17 | 4.62 | .032 |
| WAGES vs. Group-activity | 1.81 | 0.88 | 4.28 | .039 |
| Info-only vs. Group-activity | -0.07 | 1.06 | 0.00 | .952 |

**Figure 1.** *Anger reactions by intervention condition and sexism type*



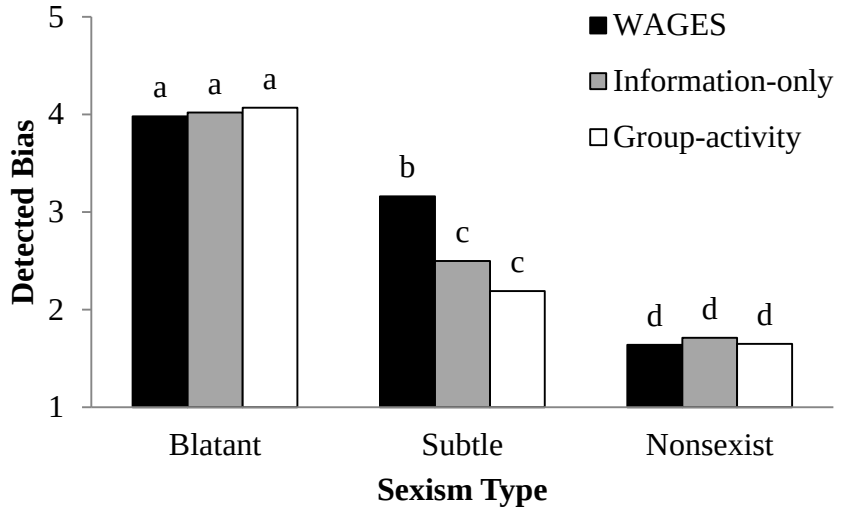

*Note*. Bars with different letters significantly differed, $p < .05$.

**Figure 2.** *Perceived unfairness by intervention condition and sexism type*



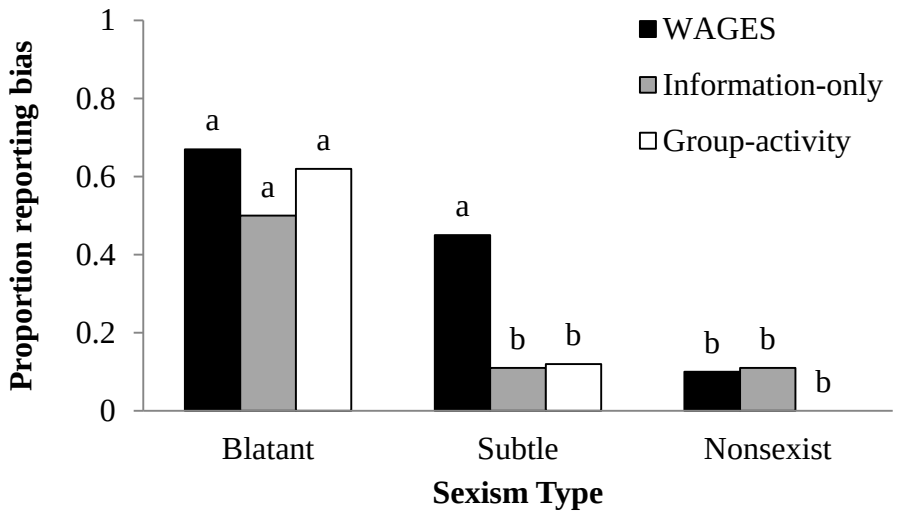

*Note*. Bars with different letters significantly differed, $p < .05$.

**Figure 3.** *Detected bias (explicit ratings) by intervention condition and sexism type*




*Note*. Bars with different letters significantly differed at $p < .05$, with one exception: in the group activity condition, the subtle sexism and nonsexist conditions differed at $p = .053$.

**Figure 4.** *Proportion of participants that reported bias within each condition*




*Note*. Bars with different letters significantly differed, $p < .05$.