Exposure to Sexism Can Decrease Implicit Gender Stereotype Bias

Miguel R. Ramos

University of Oxford and Lisbon University Institute (CIS/ISCTE-IUL)

Manuela Barreto

University of Exeter and Lisbon University Institute (CIS/ISCTE-IUL)

Naomi Ellemers

University of Leiden

Miguel Moya

University of Granada

Lúcia Ferreira

Lisbon University Institute (CIS/ISCTE-IUL)

Jimmy Calanchini

University of California, Davis

Author's note:

Please address correspondence to Miguel Ribeiro Ramos, Department of Experimental Psychology, University of Oxford, Tinbergen Building, South Parks Road, OX1 3PS, Oxford, UK. Email: miguel.ramos@psy.ox.ac.uk. This research was supported by a grant from the Portuguese Foundation for Science and Technology (PTDC/PSI-PSO/120015/2010) awarded to the second author. The first author was at Lisbon University Institute (ISCTE-IUL) when the work was carried out.

Abstract

Two studies examined the effect of exposure to sexism on implicit gender bias, focusing specifically on stereotypes of men as competent and women as warm. Male and female participants were exposed to sexism or no-sexism. In both Experiment 1 (Implicit Association Task; *N* = 115) and Experiment 2 (Go/No-go Association Task; *N* = 167), women who had been exposed to sexist beliefs demonstrated less implicit gender stereotype bias relative to women who were not exposed to sexism. In contrast, exposure to sexism did not influence men's implicit gender stereotype bias. In Experiment 2, process modelling revealed that women's reduction in bias in response to sexism was related to increased accuracy-orientation and a tendency to make warmth versus competence judgments. The implications of these findings for current understandings of sexism and its effects on gender stereotypes are discussed.

Keywords: Sexism; stereotyping; implicit bias

Word count: 137

Exposure to Sexism Can Decrease Implicit Gender Stereotype Bias

Gender stereotypes remain prevalent, among both women and men, across a variety of contexts (for reviews see Rudman & Glick, 2008; Swim & Hyers, 2009). Such stereotypes are problematic: not only do they influence how men and women are treated, but they also affect both men's and women's wellbeing and performance in gender stereotypic domains (for reviews see Barreto, 2014; Schmader, Hall, & Croft, 2014). Previous work examining the impact of sexism on gender stereotypes has mainly investigated how sexism influences women's *explicit* endorsement of gender stereotypes (e.g., Ellemers & Barreto, 2009; Kray, Thompson, & Galinsky, 2001). Our aim in this paper is to extend past work by examining the effect of exposure to sexism on men's and women's *implicit* gender stereotypic associations.

The Malleability of Gender Stereotypes

Early theorizing on implicit bias proposed that it results from automatic processes that are difficult to change (Bargh, 1999; Devine, 1989). Subsequent research, however, provided evidence that implicit stereotype bias can be quite malleable (e.g., Blair, 2002; Dasgupta & Greenwald, 2001; Lenton, Bruder, & Sedikides, 2009; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000; Rudman, Ashmore, & Gary, 2001; Wittenbrink, Judd, & Park, 2001). For example, Blair and Banaji (1996) demonstrated that implicit gender stereotype bias was affected by whether male and female names were preceded by gender consistent or inconsistent primes. Building on these findings, researchers have identified various processes underlying malleability in implicit evaluations (for reviews see Blair, 2002; Dasgupta, 2009; Gawronski & Bodenhausen, 2006; Gawronski & Sritharan, 2010; Lai, Hoffman, & Nosek, 2013; Sritharan & Gawronski, 2010). For example, Dasgupta and Rivera (2008) showed that gay bias was more malleable among individuals who had less (vs. more) contact with gay individuals, suggesting that malleability in implicit bias can be stronger when attitudes are less elaborated (see also Dasgupta & Asgari, 2004). Research has also shown that implicit racial bias may change due to shifts in emotional states (Dasgupta, DeSteno, Williams, & Hunsinger, 2009; DeSteno, Dasgupta, Bartlett, & Cajdric, 2004), exposure to counter stereotypical exemplars (e.g., Dasgupta & Greenwald, 2001), and the activation of egalitarian goals (Legault, Gutsell, & Inzlicht, 2011; Mann & Kawakami, 2012). Other contextual factors that have been shown to affect implicit biases are experimental task instructions (Blair, Ma, & Lenton, 2001), characteristics of the evaluated targets (Barden, Maddux, Petty, & Brewer, 2004), egalitarian norms (Moskowitz, Wasel, Gollwitzer, & Schaal, 1999), and the attitudes of others present in the experimental context (Lowery, Hardin, & Sinclair, 2001).

Some of these variations in implicit bias have been attributed to the contextual salience of specific identities or stereotypes, whereas others have been attributed to motivational factors (e.g., Amodio, Devine, & Harmon-Jones, 2008; Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002). Importantly, recent research has clarified that cognitive and motivational factors often work in tandem, since motivational processes can modify how information about targets is processed (Van Nunspeet, Ellemers, Derks, & Nieuwenhuis, 2014).

In addition, recent research has clarified that measures of implicit bias reflect the influence of both relatively automatic and relatively controlled processes (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; for a review see Calanchini & Sherman, 2013). That is, these measures capture both biased stereotypic associations, which are activated relatively unintentionally, and processes which constrain the expression of these biased associations (see Fiedler & Bluemke, 2005). Thus, responses to implicit measures can be subject to control, and therefore vary across contexts that affect individual motivation to control bias (e.g., Klauer & Teige-Mocigemba, 2007; Tiege-Mocigemba & Klauer, 2008).

The Present Research

Although research has shed light on some of the conditions that affect implicit bias, whether and how exposure to sexism affects implicit bias remains unexamined. Our goal in this research was to examine whether exposure to sexism would provide participants with sufficient motivation to reduce implicit gender stereotype bias and thereby disprove sexism views.

Although the effect of exposure to sexism on implicit gender stereotype bias has yet to be examined, prior research has made clear that women are often motivated to *explicitly* disconfirm gender stereotypes when these are made salient (e.g., Cihangir, Barreto, & Ellemers, 2010; Ellemers & Barreto, 2009; Kray et al., 2001). To examine whether this can also happen at the *implicit* level, we exposed participants to sexist beliefs (vs. not) and then measured their implicit gender stereotype bias. If exposure to sexism motivates individuals to be less biased, we would expect that participants who are exposed to sexism would display less implicit bias than participants who are not exposed to sexism—despite the fact that gender stereotypes are likely to be more (not less) salient in these conditions. Note, therefore, that this pattern cannot be explained without reference to motivational processes, since exposure to sexism should actually increase the cognitive salience of gender stereotypes, resulting in more bias when participants are exposed to sexism than when they are not. The argument here is, therefore, that exposure to sexism might reduce implicit bias by providing individuals with sufficient motivation to disprove sexist stereotypes.

Although motivations to disprove sexist stereotypes have not been examined in past research, prior research has investigated how people respond when they are exposed to information about traditional gender roles. This has revealed that such descriptive gender normative information can have a variety of effects: In some cases it can increase gender stereotype bias (Rudman & Phelan, 2010) and induce stereotype threat (Davies, Spencer, Quinn, & Gerhardstein, 2002), but in other cases it can also decrease gender stereotype bias (e.g., de Lemus, Spears, Bukowski, Moya, & Lupiáñez, 2013). Importantly, these prior investigations primed gender roles in a purely descriptive manner (i.e., with TV commercials or photographs of men and women in stereotypical occupations). Gender role depictions are, however, ambiguous: They can be interpreted as communicating the appropriateness of gender roles, but they can also be interpreted as caricatures and expressions of disapproval. Sexist statements are different: They provide clarity about people's beliefs, and are strongly prescriptive, in that they communicate the conviction that men and women should *comply* with traditional gender roles (e.g., Burgess & Borgida, 1999; Fiske & Stevens, 1993). As such, the results of prior work are not entirely relevant to understand or predict how implicit gender stereotype bias is affected when people are exposed to sexist views, which is the focus of the current research. In addition, previous work did not examine the processes through which exposure to sexism affects stereotype associations. The present research therefore also extends previous knowledge by using process modelling to understand how exposure to sexist beliefs may impact on gender stereotypical associations.

To examine our hypothesis, we measured participants' implicit gender stereotype bias after exposure to sexist beliefs and compared this to when participants were not exposed to sexism. In addition, while prior research examined gender stereotypical associations as a function of exposure to traditional gender roles in comparison to control conditions, we take on board knowledge that sexism is expressed in multiple ways and compare the effects of two types of sexism. We therefore included two experimental conditions: Exposure to hostile sexism and exposure to benevolent sexism (Glick & Fiske, 1996; 2001). Because both forms of sexism are based on the same gender stereotype of women as warm but incompetent, we did not expect to find differences across these two sexism conditions, but this method allowed us to test for this possibility. In addition, we separately examined responses from female and male participants. Although women are more likely to be motivated than men to reject sexist views, it is possible that similar motives, as well as others, drive men to also reduce their gender stereotype bias when exposed to sexism (e.g., heightened vigilance so as not to appear biased, Devine et al., 2002; Monteith, Lybarger, & Woodcock, 2009; Van Nunspeet et al., 2014). Given the dearth of research on men's responses to sexism targeting women, we did not make specific predictions regarding men's implicit gender stereotype bias.

In two experiments, we exposed male and female participants to sexism (either hostile or benevolent sexist beliefs) or no sexism. In Experiment 1, we measured gender stereotype bias with an Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1988). In Experiment 2, we sought a more nuanced understanding of the pattern of gender stereotype bias by using a Go/No-go Association Task (GNAT; Nosek & Banaji, 2001). In both experiments, we predicted that female participants who were exposed to sexism would display less gender stereotype bias than women who were not exposed to sexism. We had no specific predictions for male participants.

Experiment 1

Method

Design and participants. A total of 55 males and 60 females (age 17-35, M=21.29, SD=3.18) were randomly and equally assigned to 3 exposure conditions: no sexism vs. benevolent sexism (BS) vs. hostile sexism (HS).

Procedure. University students were invited to the laboratory where they were asked to perform a memory task. They observed, memorized, and later recalled the association between six photos and sentences (procedure adapted from Dasgupta & Asgari, 2004; Study 1). The photos portrayed men and women interacting with each other and the sentences were derived from Glick and Fiske's (1996) Ambivalent Sexism Inventory (in the sexism conditions) or created for this study (no sexism condition; please see Appendix for the full wording of the items used). In the BS condition, participants read sentences from the benevolent sexism subscale (e.g., "No matter how accomplished he is, a man is not truly complete as a person unless he has the love of a woman"); in the HS condition, participants read sentences from the hostile sexism subscale (e.g., "Most women interpret innocent remarks or acts as being sexist"). In the no sexism condition, participants saw photos of flowers and a short description of their names and origin (e.g., "Water lilies are aquatic plants that can often cut the flow of water").

After this task, participants were invited to participate in an ostensibly separate study involving the assessment of learning abilities. This corresponded to an IAT with 7 blocks (Greenwald, Nosek, & Banaji, 2003) with male and female names as target categories, and words related to competence and warmth as attribute categories. Gender stereotypes are structured around the perception that men are particularly competent whereas women are predominantly warm (Eagly & Mladinic 1989; Eagly & Steffen 1984; Fiske et al. 2002; Langford & MacKinnon, 2000). Thus, this IAT allows us to examine the extent to which participants hold implicit gender stereotype bias. (Ebert, Steffens, & Kroth, 2014). In the IAT's congruent blocks, participants were instructed to use one response key for male names and words related to competence, and another response key for female names and words related to warmth. In incongruent blocks, this pattern was reversed such that female names were paired with competence attributes and male names with warmth attributes. Practice blocks (1, 2, 3, 5, and 6) consisted of

20 trials each, whilst test blocks (4 and 7) included 40 trials per block. Each stimulus was presented on the computer screen until participants provided their response. There was a 150ms interval between each stimulus. Within each block, stimuli were randomly selected without replacement and no more than two consecutively presented stimuli belonged to the same category (Greenwald et al., 1998). Names were presented in white font, whilst stereotypical attributes were presented in green font (against a black background). The order of congruent and incongruent blocks was counterbalanced.

Stimuli for the IAT were selected according to two pilot studies examining their typicality (as indicators of competence and warmth) and valence. In the typicality pilot study, 42 participants (20 females) rated the extent to which each attribute (from a list of 220 attributes commonly used in stereotype research) was seen in Portuguese society as typically masculine or feminine. All attributes were rated in 7-point scales, from (1) “typically masculine” to (7) “typically feminine” with the midpoint labelled as “neither masculine nor feminine.” In the subsequent valence pilot study, 40 participants (23 females) were presented with the attributes that had been judged as most stereotypical and rated the valence of each attribute. The 8 attributes that were considered most typically feminine were: Affectionate, gentle, considerate, understanding, kind, friendly, good, and warm (in the original Portuguese materials these were: Afectuoso, meigo, atencioso, compreensivo, afável, amável, bondoso, and carinhoso). The 8 most typically masculine attributes were: Confident, practical, competitive, active, determined, objective, reliable, and leader (original Portuguese: Confiante, prático, competitivo, activo, determinado, objectivo, seguro, and líder). Stereotypical male attributes were rated as slightly less positive than stereotypical female attributes among females and slightly more positive among males. Although not ideal, this difference in valence is common in gender stereotype research (Rudman, Greenwald, & McGhee, 2001) and was deemed unproblematic since our focus was to examine stereotype associations across conditions. As also observed in past research, while female stereotypical attributes generally tapped onto the warmth dimension, male stereotypical attributes reflected competence. Accordingly, the attributes included in the IAT were labelled as *warmth* and *competence*. The target stimuli included in the IAT comprised male and female names, in Portuguese, that were of similar size and familiarity (e.g., male names: João, Paulo, and Mário; female names: Sara, Joana, and Maria).

Results and Discussion

Data screening. In line with standard guidelines to analyse IAT data, (Greenwald et al., 2003), we eliminated all trials with latencies greater than 10,000ms (there were 22 trials) and data from all participants for whom more than 10% of trials had response latencies lower than 300ms (there were 3 participants in this situation) . To calculate *D* scores, we computed the difference in reaction times on incongruent and congruent trials divided by a pooled standard deviation of all trials for each participant. Higher *D* scores thus represent stronger implicit gender stereotype bias.

IAT analysis. *D* scores were analysed with an ANOVA with participants' gender and sexism exposure as between-participant factors. This revealed no reliable main effects of participants' gender, $F(1,109)=0.25$, $p=.616$, $\eta p^2=.002$, or of sexism exposure, $F(2,109)=0.21$, $p=.822$, $\eta p^2=.004$. However, a reliable interaction emerged, $F(2,109)=3.41$, $p=.037$, $\eta p^2=.059$ (see Figure 1). To test our predictions, we conducted planned comparisons between the no sexism condition (weight = 2) and the two sexism conditions (weights of -1 for each BS and HS condition). These analyses showed that women had weaker gender stereotype bias when they were exposed to sexist beliefs ($M_{BS}=0.27$; $SD=0.36$ and $M_{HS}=0.27$; $SD=0.39$) than when they were not ($M=0.53$; $SD=0.30$), $F(1,109)=5.21$, $p=.024$, $\eta p^2=.046$. For men, there were no differences between the no sexism (M=0.28; SD=0.47) and the sexism conditions (M_{BS}=0.50; SD=0.49 and M_{HS}=0.42; SD=0.45), $F(1,109)=2.06$, $p=.154$, $\eta p^2=.019$.[1]

In sum, consistent with predictions, women showed weaker gender stereotype bias when exposed to benevolent or hostile sexist beliefs than following no exposure to sexism. Gender stereotype bias among men did not vary across conditions. This pattern of results supports our hypothesis that women reduce stereotypic gender bias when they are exposed to sexism as a response against sexist views.

Experiment 2

Experiment 2 aims to further specify how exposure to sexism reduces women's implicit gender stereotype bias by examining whether it strengthens the relationship between women and competence, weakens the relationship between women and warmth, or both. To do so, we used a Go/No-go Association Task (GNAT; Nosek & Banaji, 2001), which allows us to make orthogonal comparisons on the dimensions of interest (i.e., male-competence, male-warmth, female-competence, female-warmth), rather than relying on the inherently-relative IAT *D* score. Also, we modified the no sexism condition from Study 1 to exclude the possibility that flowers, which are more stereotypically related to women than men, had influenced responses in the no sexism condition. As such, the stimuli for the no sexism condition in Experiment 2 consisted of photos of keys to different locks.

To shed light on the processes underlying participants' responses to sexism, we drew on the quadruple process model (Quad model: Conrey et al., 2005; Sherman et al., 2008), which is designed to estimate the contribution of four qualitatively distinct processes to responses on implicit measures of bias (for a review, see Calanchini & Sherman, 2013)[2]. The four processes specified by the Quad model are the activation of biased associations (association activation; AC), the ability to determine the correct response (detection; D), the ability to overcome biased associations when they conflict with the correct response (overcoming bias; OB), and other types of bias that guide responses in the absence of other available guides to responding (guessing; G). AC represents mental associations (e.g., stereotypes, attitudes) activated by the specific stimuli of a given implicit measures. In contrast, D and OB reflect domain-general accuracy-oriented and inhibitory processes, respectively, which constrain activated associations from influencing responses (Calanchini, Sherman, Klauer, & Lai, 2014). Finally, the G parameter, which represents guessing biases that drive responses when neither AC nor D is activated, is not necessarily random but, instead, may be quite strategic, such as a general positivity bias or a preference for stimuli on one side of the display (Conrey et al., 2005; Nisbett & Wilson, 1977). In line with our argument, we expected exposure to sexism would motivate women to disconfirm gender stereotypes. Given that mental associations such as stereotypes are assumed to be activated unintentionally upon exposure to a relevant stimulus (Bargh, 1999), we expected that reactions against sexism would influence one of the non-associative processes that contribute to responses—D, OB, or G—although we had no firm predictions regarding which of these processes would be influenced.

Method

Design and participants. A total of 76 male and 92 female participants (age 18-31, *M*=20.50, *SD*=2.37) were randomly assigned to one of 3 conditions: no sexism vs. benevolent sexism vs. hostile sexism.

Procedure. The procedure was the same as in Experiment 1, but in this study the no sexism condition consisted of six pictures of keys to different locks accompanied by related sentences (e.g., “Despite not revealing to whom it belongs, each key always has a story to tell”). Participants responded to a GNAT using the same attributes as in

Experiment 1. Competence and warmth attributes appeared twice as frequently as male and female names (see Nosek & Banaji, 2001). Participants were asked to indicate the presence of a specific target concept or pairs of concepts with a keyboard press, and to withhold responses to all other stimuli. The GNAT included 6 blocks and followed Nosek and Banaji's guidelines (2001): (1) initial target-concept discrimination (i.e., pressing a key only if male names are shown; pressing a key only if female names are shown), (2) attribute discrimination (i.e., pressing a key only if warmth attributes are shown; pressing a key only if competence attributes are shown), (3) first combined task (female counter-stereotypical: pressing a key only if female names or competence attributes are shown), (4) second combined task (female stereotypical: pressing a key only of male names or warmth attributes are shown), (5) third combined task (male counter-stereotypical: pressing a key only if male names or warmth attributes are shown), and (6) fourth combined task (male stereotypical: pressing a key only if male attributes or competence attributes are shown). Half of the participants followed this block order, whilst the other half responded to the stereotypical block first and then the counter-stereotypical block. Practice blocks (1 and 2) consisted of 20 trials each, whilst test blocks (3, 4, 5, and 6) comprised 20 practice trials and 50 test trials (in each block). In all blocks, the distracter items were items from a non-matching concept (e.g., a competence attribute when participants were prompted to press a key every time they saw a warmth attribute). Following the procedures recommended by Nosek and Banaji (2001), participants responded to the six blocks with a response deadline of 1000 milliseconds (ms) and then again to the same blocks with a faster response deadline of 833ms. The response deadlines for target trials was twice that for distracter items such that participants had a 1000ms and 833ms for the distracter items and a 2000ms and

1666ms deadline for the target trials. In all trials there was a 150ms interstimulus interval.

Note that the response deadlines chosen for this study correspond to those recommended by Nosek and Banaji (2001) as appropriate for analyses of response times. Nosek and Banaji (2001; Experiment 5) concluded that examining sensitivity (i.e., based on response accuracy) or reaction times produces identical results. We elected to first analyze response times in order to maintain consistency and enable comparison with Experiment 1 (where an IAT with response times was used), but we also utilized response accuracy for the process analyses.

Results and Discussion

Data screening. In line with recommendations to analyse the GNAT (Nosek & Banaji, 2001), response latencies lower than 200 milliseconds were removed (there were 465 trials in this situation) and data from participants who had an error rate greater than 40% were deleted (there was only one participant with a 40% error rate).

GNAT analysis. All responses to distracter trials were excluded and the analyses reported correspond to the reaction times of the correct responses in “go” trials. Analyses of response latencies followed a 3 (exposure: no sexism vs. BS beliefs vs. HS beliefs) X 2 (participant gender: male vs. female) X 2 (target gender: male vs. female) X 2 (dimension: competence vs. warmth) ANOVA with repeated measures on the last two factors. Results revealed a reliable main effect of dimension, $F(1,161)=20.95$, $p<.001$, $\eta p^2=.061$, and 2-way interactions between dimension and target gender, $F(2,161)=12.82$, $p<.001$, $\eta p^2=.038$, and between dimension and sexism exposure, $F(2,161)=5.48$, $p=.005$, $\eta p^2=.033$. There was also a 4-way interaction between sexism exposure, participant gender, target gender, and dimension, $F(2,161)=5.13$, $p=.006$, ηp2=.031. We decomposed the 4-way interaction by examining lower order effects per dimension.

Effects on competence. A 3 (sexism exposure: no sexism vs. BS vs. HS) X 2 (participant gender: male vs. female) X 2 (target gender: male vs. female) was conducted on response times for competence judgments. Results showed a reliable main effect of target gender, $F(1,161)=5.29$, $p=.022$, $\eta p^2=.016$, and a reliable interaction between participant gender and sexism exposure, $F(2,161)=7.91$, $p<.001$, $\eta p^2=.047$. Both effects were qualified by a reliable 3-way interaction between participant gender, target gender, and sexism exposure, $F(2,161)=4.17$, $p=.016$, $\eta p^2=.025$ (see Figure 2). As in Experiment 1, the 3-way interaction was first examined for women and then for men.

To test our hypotheses we performed planned comparisons contrasting the no sexism condition (weight = 2) with the two sexism conditions (weighted as -1 each). Consistent with Experiment 1, t-tests showed that when women were exposed to sexist beliefs, they displayed less stereotype bias. Specifically, women responded more quickly when female names and competence words shared a response key (M_{BS}=736ms; SD=105.83 and M_{HS}=784ms; SD=101.41) than when they were not exposed to sexism (*M*=818ms; *SD*=118.46), $t(92)=5.93$, $p=.017$, $\eta p^2=.062$. Additionally, when women were not exposed to sexist beliefs, they responded more quickly when competence words shared a response key with male names (*M*=745ms, *SD*=108.94) than with female names (*M*=818ms, *SD*=118.46) $F(1,161)=5.13$, $p=.024$, $\eta p^2=.016$.[3]). However, when exposed to sexist beliefs, women responded equally quickly when competence words shared a response key with male and female names, $t(161)=0.35$, p=.852, $\eta p^2<.001$.

For male participants, planned comparisons contrasting the no sexism condition (M=740ms; SD=128.15) against the sexism conditions (M_{BS}=780ms; SD=123.00 and

M_{HS}=707ms; SD=127.95) revealed no reliable effects, $F(1,161)=0.22$, $p=.883$, $\eta p^2<.001$.

Effects on warmth. The same analyses were performed on responses to warmth trials. No reliable effects emerged: Main effects of sexism exposure, $F(2,161)=0.53$, $p=.589$, $\eta p^2=.003$, target gender, $F(1,161)=0.88$, $p=.350$, $\eta p^2=.003$, and participant gender, $F(1,161)=0.04$, $p=.837$, $\eta p^2<.001$. Interactions between sexism exposure and target gender, $F(2,161)=0.08$, $p=.928$, $\eta p^2<.001$, between target gender and participant gender, $F(1,161)=0.12$, $p=.731$, $\eta p^2<.001$, and between sexism exposure, target gender, and participant gender $F(2,161)=0.46$, $p=.630$, $\eta p^2=.003$ (see Figure 3).

Process analysis. We applied the Quad model to these data according to Sherman et al., (2008) by estimating parameters for AC, D, OB, and G based on the number of correct and incorrect responses on the GNAT. Two separate AC parameters were estimated: One measuring the extent to which associations between female and warmth (FAC) were activated in performing the task and another measuring the extent to which associations between male and competence (MAC) were activated. The G parameter is anchored at .5 (i.e., no bias), and was coded so that scores greater than .5 represent a tendency for responses that favour competence over warmth (i.e., a “go” response on a trial which included competence as target category; a “no go” response on a trial which included warmth as a target category) and scores less than .5 represent a tendency for responses that favour warmth over competence (i.e., a “go” response on a trial which included warmth as target category; a “no go” response on a trial which included competence as a target category). Given that we expected different results for women and men, we divided the data by participant gender and performed separate analyses (see Table 1 for parameter estimates).

The overall error rate for female participants on the GNAT was 8.2% and the Quad model fit the data well, $\chi^2(df=1)=10.15$, $p=.12$, $w=0.03$. Female participants' accuracy-oriented detection (D) increased following exposure to sexism ($M=0.87$) relative to the no sexism condition ($M=0.80$), $\Delta\chi^2=37.61$, $p<.001$, $w=0.05$. Additionally, exposure to sexism biased female participants' responses (G) towards warmth judgments and away from competence judgments ($M=0.45$) relative to the no sexism condition ($M=0.52$), $\Delta\chi^2=4.16$, $p=.04$, $w=0.02$. However, exposure to sexism had no influence on female participants' activation of female-warmth (FAC) or male-competence (MAC) stereotypic associations, or their ability to overcome those biased associations (OB), all $\Delta\chi^2$s<1.34, *p*s>.24, *w*s<0.01.

The overall error rate of male participants on the GNAT was 6.9% and the Quad model fit the data well, $\chi^2(df=1)=12.33$, $p=.06$, $w=0.03$[4]. None of male participants' Quad parameter estimates differed across sexism and control conditions, all $\Delta\chi^2$s<2.43, *p*s>.11, *w*s≤0.01.[5]

The results of Experiment 2 replicate those of Experiment 1: Women exposed to sexist beliefs subsequently demonstrated decreased implicit gender stereotype bias, but men did not. Experiment 2 also provides insight into why women's bias decreases in response to sexism. When exposed to sexism versus not exposed, women respond more quickly when female names are paired with words related to competence, but do not change their responses to words related to warmth. As such, it appears that women do not reject stereotypes favourable to women (warmth), but assert women's qualities in the dimension in which they are stereotypically portrayed as inferior to men (competence). This study also enables us to specify the process underlying this pattern of responses. Indeed, application of the Quad model revealed that exposure to sexism increases women's accuracy-oriented detection and, in the absence of any other guides to response, biases their responses towards warmth judgments and away from competence judgments.

General Discussion

Across two experiments, we investigated whether exposure to sexism can reduce implicit gender stereotype bias. When exposed to sexist beliefs, women but not men showed decreased implicit gender stereotype bias relative to a no sexism condition. Moreover, exposing women to sexism increased the speed with which they responded to female/competence pairings, increased their accuracy orientation, and biased their responses in a way that favoured warmth over competence.

Our findings build on prior research demonstrating the malleability of implicit stereotype bias (e.g., Blair & Banaji, 1996; Blair et al., 2001), as well as research showing that women are motivated to disconfirm explicit gender stereotypes (e.g., Becker & Wright, 2011; de Lemus et al., 2013; Ellemers & Barreto, 2009; Spears, Jetten, & Doosje, 2001). However, prior research had not yet examined whether exposure to sexism was sufficient to motivate reduced implicit gender stereotype bias. We therefore extend prior work by additionally showing that women may modify their implicit gender stereotype bias when exposed to different types of sexist statements.

It is important to note that this pattern of results can only be explained by reference to individual motivation. A priori, we expected that women and men would react to sexist statements differently, given that men do not suffer the same deleterious consequences of sexism as women and, therefore, do not have the same motivation to disconfirm gender stereotypes. Moreover, these results support a motivational explanation because a cognitive salience explanation would lead to the opposite pattern of results. Indeed, if sexist statements increased the cognitive salience of gender stereotypes, then we would expect greater rather than less implicit gender bias.

However, women who were exposed to sexism actually demonstrated less implicit gender bias. Process modelling supports these conclusions by demonstrating that exposure to sexism did not increase the activation of stereotypic associations (FAC, MAC) but, instead, increased an accuracy-oriented control process (D) and changed women's response tendencies (G). The effects we found with the D parameter among females were relatively large, which further suggests that the effect of sexism on bias was primarily driven by increasing accuracy orientated detection (D). In other words, reduced gender bias following exposure to sexism appears not to be related to changes in the activation of biased stereotypic associations but, rather, is the result of other non-associative processes that drive responses. Detection is a domain-general process that is associated with activity in brain areas which are generally associated with control and can reflect how motivated a person is to perform well on a given task (Beer et al., 2008; Calanchini et al., 2014; Sherman et al., 2008). In contrast, Guessing is a domain-specific process that guides responses when biased associations (i.e., stereotypes) are not activated and the correct answer is not detected (Calanchini et al., 2014). The curious reader may wonder why exposure to sexism did not also increase women's ability to overcome biased associations (OB). One possible explanation is mathematical: the OB parameter is estimated from fewer trials in the Quad model than the AC, D, and G parameters, and this relative lack of reliability sometimes makes it challenging to detect between-person or -group differences on OB (though this is not always the case; for a review, see Calanchini & Sherman, 2013). Another possible explanation for why exposure to sexism did not increase women's OB is because exposure to sexism did not increase the activation of female-competence or male-warmth stereotypes. Given that sexism did not make biased associations more salient, perhaps there was simply no need to overcome them.

Taken together, the present research contributes to existing knowledge by clarifying the conditions under which individuals who are targeted by gender stereotypes are likely to react against them, and the mechanisms by which they might do so. We expected and showed that women are less likely to react against gender stereotypes when they have not been recently exposed to sexist beliefs. Perhaps ironically, this suggests that women who are chronically exposed to sexism are more motivated to disconfirm gender stereotypes than women who do not regularly encounter sexism.

Prior work has suggested that changing gender stereotypes is particularly hard because such stereotypes involve a trade-off for women, who can be perceived as *either* competent *or* warm (Cuddy, Fiske, & Glick, 2004). However, Experiment 2 indicates that this is not always necessarily the case: our female participants adjusted their responses when women were associated with competence, but maintained their responses when women were associated with warmth. This further underscores the motivational, and perhaps even strategic, nature of these responses. That is, women reduced implicit gender stereotype bias by changing how they respond when their gender is associated with a dimension on which they are seen as stereotypically inferior (competence) without yielding their superiority on a dimension that favors them (warmth). This resonates with past research conducted with explicit measures. Even though women often seek to be perceived as competent because with competence comes status, power, and resources (Fiske et al., 2002; see for reviews Fiske, 2010; Fiske, Cuddy, & Glick, 2007), they also tend to associate themselves with warmth (Rudman, Greenwald, & McGhee, 2001) and are often reluctant to shed perceptions of warmth (Rudman & Kilianski, 2000).

Challenging sexism by adjusting implicit gender stereotype bias has important implications for women in a variety of contexts. For example, research has shown that exposure to subtle gender stereotypes can impair women's performance on tasks in which they are deemed stereotypically inferior (Spencer, Steele, & Quinn, 1999). Our findings suggest that this is not inevitable since women can challenge even stereotypes to which they are exposed in subtle ways. Also, women can do so by adjusting their implicit biases, which in turn have been closely linked to task performance (Levy, 2003), placing them in a good position to disprove gender stereotypes.

Though women were influenced by exposure to sexism, sexism had no influence on men's implicit gender stereotype bias. One reason for this apparent null effect might be that men, as the dominant group, are less motivated to disprove gender stereotypes that favor them. Because it does not target them directly, men are less likely to be vigilant about sexism and the contextual cues presented in our studies. Another possible explanation for men's lack of response to exposure to sexism is that, as the dominant group, men may be generally more chronically concerned with being seen as prejudiced (Vorauer & Kumhyr, 2001; Vorauer, Hunter, Kelley, & Scott, 2000) and, therefore, may learn to control their biased responses across a variety of contexts. Even though men's responses were relatively biased across all conditions, they may have still been as controlled as possible, leaving little opportunity for further monitoring when they were exposed to sexist beliefs. Future research might further distinguish between these possibilities.

Finally, in these studies, women responded similarly when exposed to benevolent sexism and hostile sexism. This is not surprising, given that benevolent and hostile sexism overlap substantially, correlate positively across gender and cultural groups, and work in tandem to communicate the same ideal view of women (Glick &

Fiske, 1996; 2001; Glick et al., 2000). Nevertheless, the fact that women demonstrated reduced implicit gender bias in response to benevolent sexism extends previous research, which has shown that women have difficulty objecting to benevolent sexism (Barreto & Ellemers, 2005; Becker & Wright, 2011). Specifically, our data suggests that despite the prior finding that it is hard to object to benevolent (versus hostile) sexism, women are able to challenge these beliefs by adapting (i.e., reducing) their stereotype bias. Indeed, the fact that exposure to BS reduces women's self-confidence and can impair their task performance (e.g., Dardenne et al., 2007) suggests that women might be motivated to also dispute BS. Our research shows that this happens at least at the level of stereotype associations.

This research demonstrates that while women can demonstrate gender stereotype bias, they are less likely to do so when reminded of sexist beliefs. Understanding how sexism might affect gender stereotype bias has potential implications for prejudice reduction interventions. Future research might investigate whether additional information, e.g., about the pervasiveness of sexism, affects implicit and explicit stereotyping. It would also be fruitful for future research to examine how the observed changes in gender stereotype bias following from sexism exposure may further motivate behavioural changes and how lasting these effects might be. Regardless, this research has taken us a step further in understanding contexts where, and ways in which, women resist sexism.

References

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology, 94,* 60–74.

Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M.B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology, 87*, 5–22.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York: Guilford.

Barreto, M. (2014). Experiencing and coping with stigma. In J. A. Simpson , J. F. Dovidio (Ed). *APA handbook of personality and social psychology, Volume 2: Group processes. APA handbooks in psychology*, (pp. 473-506). Washington, DC, US: American Psychological Association. doi: 10.1037/14342-018

Barreto, M., & Ellemers, N. (2005). The burden of benevolent sexism: how it contributes to the maintenance of gender inequalities. *European Journal of Social Psychology, 35,* 633–642.

Becker, J., & Wright, S. C. (2011). Yet another dark side of chivalry: Benevolent sexism undermines and hostile sexism motivates collective action for social change. *Journal of Personality and Social Psychology, 101*, 62–77

Beer, J. S., Stallen, M., Lombardo, M. V., Gonsalkorale, K., Cunningham, W. A., & Sherman, J. W. (2008). The Quadruple Process model approach to examining the neural underpinnings of prejudice. *Neuroimage, 43,* 775-783.

Blair, I.V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*, 242–61.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology, 70*, 1142-1163.

Blair, I., Ma, J., & Lenton, A. (2001). Imagining stereotypes away: The moderation of automatic stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*, 828–841. doi 10.1037/0022-3514.81.5.828

Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive stereotyping in sex discrimination. *Psychology, Public Policy, and Law, 5*, 665–692.

Calanchini, J., & Sherman, J. W. (2013). Implicit attitudes reflect associative, non-associative, and non-attitudinal processes. *Social and Personality Psychology Compass*, *7*, 654-667.

Calanchini, J., Sherman, J. W., Klauer, K. C., & Lai, C. K. (2014). Attitudinal and non-attitudinal components of IAT performance. *Personality and Social Psychology Bulletin*, *40*, 1285-1296.

Cihangir. S., Barreto B., & Ellemers, N. (2010). The dark side of ambiguous discrimination: How state self-esteem moderates emotional and behavioural responses to ambiguous and unambiguous discrimination. *British Journal of Social Psychology, 49*, 155–174.

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating multiple processes in implicit social cognition: the quad model of implicit task performance. *Journal of Personality and Social Psychology*, *89*, 469.

Cuddy, A.J.C., Fiske, S.T., & Glick, P. (2004). When professionals become mothers, warmth doesn’t cut the ice. *Journal of Social Issues, 60*, 701-718.

Dardenne, B., Dumont, M., & Bollier, T. (2007). Insidious dangers of benevolent sexism: consequences for women's performance. *Journal of Personality and Social Psychology, 93*, 764–779. doi:10.1037/0022-3514.93.5.764

Dasgupta, N. (2009). Mechanisms underlying malleability of implicit prejudice and stereotypes: The role of automaticity versus cognitive control. In T. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination*. Mahwah, NJ: Erlbaum.

Dasgupta, N., Greenwald, A.G. (2001). On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81*, 800–14.

Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*, 642–658. Doi 10.1016/j.jesp.2004.02.003

Dasgupta, N., & Rivera, L. M. (2008). When social context matters: The influence of long-term contact and short-term exposure to admired outgroup members on implicit attitudes and behavioral intentions. *Social Cognition, 26*, 112–123.

Dasgupta N., DeSteno D. A., Williams L., Hunsinger M. (2009). Fanning the flames of prejudice: The influence of specific incidental emotions on implicit prejudice. *Emotion, 9*, 585–591

Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin, 28*, 1615–1628.

de Lemus S., Spears R., Bukowski M., Moya M., & Lupianez J. (2013). Reversing implicit gender stereotype activation as a function of exposure to traditional gender roles. *Social Psychology*, *44*, 109–117.

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

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, *82*, 835–848.

DeSteno, D., Dasgupta, N., Bartlett, M., & Cajdrie, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science, 15*, 319 – 324

Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin, 15*, 543–558.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*, 735–754.

Ebert, I. D., Steffens, M. C., & Kroth, A. (2014). Warm, but maybe not so competent? Contemporary implicit stereotypes of women and men in Germany. *Sex Roles*, 70, 359 – 375. doi:10.1007/s11199–014-0369–5

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

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. *Basic and Applied Social Psychology*, *27*, 307–316

Fiske, S. T. (2010). Interpersonal stratification: Status, power, and subordination. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 941–982). New York, NY: Wiley.

Fiske, S.T., Cuddy, A.J.C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*, 77–83.

Fiske, S. T., Cuddy, A. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902.

Fiske, S. T., & Stevens, L. E. (1993). What's so special about sex? Gender stereotyping and discrimination. In S. Oskamp, & M. Costanzo (Eds.), *Gender issues in contemporary society* (pp. 173–196). Thousand Oaks, CA: Sage

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692 – 731.

Gawronski, B., & Sritharan, R. (2010). Formation, change, and contextualization of mental associations: Determinants and principles of variations in implicit measures. In B. Gawronski & B. K. Payne (Eds.*), Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 216–240). New York: Guilford Press.

Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology, 70*, 491–512.

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

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J., Abrams, D., Masser, B., et al. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology*, 79, 763–775.

Greenwald, A.G., McGhee, D.E., Schwartz, J.L.K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–80.

Greenwald, A.G., Nosek, B.A., Banaji, M.R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

Kawakami, K., Moll, J., Hermsen, S., Dovidio, J.F., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*, 871–888. Doi 10.1037//0022-3514.78.5.871

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology, 43*, 648 – 655.

Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology, 80*, 942-958.

Lai, C. K., Hoffman, K. M., & Nosek, B.A. (2013b) Reducing implicit prejudice. *Social and Personality Psychology Compass, 7*, 315-330.

Langford, T., & MacKinnon, N. J. (2000). The affective bases for the gendering of traits: Comparing the United States and Canada. *Social Psychology Quarterly, 63*, 34 – 48.

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

Lenton, A. P., Bruder, M., & Sedikides, C. (2009). A meta-analysis on the malleability of automatic gender stereotypes. *Psychology of Women Quarterly, 33,* 183–196. doi:10.1111/j.1471-6402. 2009.01488.x

Levy B.R. (2003). Mind matters: Cognitive and physical effects of aging self-stereotypes. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 58,* 203–211.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81,* 842–855. doi:10.1037/0022-3514.81.5.842.

Mann, N.H., & Kawakami, K. (2012). The long, steep path to equality: Progressing on egalitarian goals. *Journal of Experimental Psychology: General, 141,* 187-197

Montañés, P., de Lemus, S., Bohner, G., Megías, J. L., Moya, M., & Garcia-Retamero, R. (2012). Intergenerational transmission of benevolent sexism from mothers to daughters and its relation to daughters' academic performance and goals. *Sex Roles, 66,* 468–478. doi:10.1007/s11199-011-0116-0

Monteith, M. J., Lybarger, J. E., & Woodcock, A. (2009). Schooling the cognitive monster: The role of motivation in the regulation and control of prejudice. *Social and Personality Compass, 3,* 211–226.

Moskowitz, G. B., Wasel, W., Gollwitzer, P. M., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77,* 167—184.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Nosek, B. A., & Banaji,M. R. (2001). The Go/No-Go Association Task. *Social Cognition, 19*, 625–666.

Rudman, L. A., Ashmore, R. D., Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 81*, 856–868.

Rudman, L. A., & Glick, P. (2008). *The social psychology of gender: How power and intimacy shape gender relations*. New York, Guilford Press.

Rudman, L. A., Greenwald, A. G., & McGhee, D. E. (2001). Implicit self-concept and evaluative implicit gender stereotypes: Self and ingroup share desirable traits. *Personality and Social Psychology Bulletin, 27*, 1164–1178.

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26*, 1315–1328.

Rudman, L. A., & Phelan, J. E. (2010). The effect of priming gender roles on women's implicit gender beliefs and career aspirations. *Social Psychology, 41*, 192–202.

Schmader T, Hall W, & Croft A. (2014). Stereotype threat in intergroup relations. In J. Simpson, J. F. Dovidio (Ed). *APA handbook of personality and social psychology*. Washington DC; American Psychological Association.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115*, 314-335.

Sibley, C. G., Wilson, M., & Duckitt, J. (2007). Antecedents of men's hostile and benevolent sexism: The dual roles of social dominance orientation and right-wing authoritarianism. *Personality and Social Psychology Bulletin, 33*, 160-172.

Spears, R., Jetten, J., & Doosje, B. (2001). The (il)legitimacy of ingroup bias: From social reality to social resistance. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations* (pp. 332-362). New York: Cambridge University Press.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4 –28. doi:10.1006/jesp.1998.1373

Sritharan, R., & Gawronski, B. (2010). Changing implicit and explicit prejudice: Insights from the Associative-Propositional Evaluation Model. *Social Psychology, 41*, 113-123

Swim, J. K., & Hyers, L. L. (2009). Sexism. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping and discrimination* (pp. 407–430). New York: Psychology Press.

Teige-Mocigemba, S., & Klauer, K. C. (2008). "Automatic" evaluation? Strategic effects on affective priming. *Journal of Experimental Social Psychology, 44*, 1414 –1417

van Nunspeet, F., Ellemers, N., Derks, B., & Nieuwenhuis, S. (2014). Moral concerns increase attention and response monitoring during IAT performance: ERP evidence. *Social and Cognitive Affective Neuroscience, 9*, 141-149. doi: 10.1093/scan/nss118

Vorauer, J. D., & Kumhyr, S. M. (2001). Is this about you or me? Selfversus other-directed judgments and feelings in response to intergroup interaction. *Personality and Social Psychology Bulletin, 27*, 706–719

Vorauer, J. D., Hunter, A. J., Main, K. J., & Roy, S. A. (2000). Metastereotype activation: Evidence from indirect measures for specific evaluative concerns experienced by members of dominant groups in intergroup interaction. *Journal of Personality and Social Psychology, 78*, 690 –707.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81*, 815–827. doi:10.1037/0022-3514.81.5.815

Footnotes

1. We also performed multiple comparisons between the three different conditions. Results showed that women displayed weaker stereotype bias when exposed to BS (M=0.28; SD=0.34) than to the no sexism condition (M=0.53; SD=0.30), $F(1,109)=4.09$, $p=.046$, $\eta p^2=.036$. Women's bias was also weaker in the HS condition (M=0.27; SD=0.39), than in the no sexism condition, $F(1,109)=3.92$, $p=.050$, $\eta p^2=.035$. There were no differences between the BS and HS conditions, $F(1,109)=0.01$, $p=.961$, $\eta p^2<.001$.

2. We could not perform a similar analysis in Experiment 1 because in that study the error rates were not recorded. This does not threaten the validity of the analysis reported. Experiment 1 was designed so that participants had to provide a correct response after any error. The analyses examine the total latency required to achieve a correct response, which has been found to be an equivalent alternative to Greenwald et al.'s. (2003) improved algorithm (see Table 4, Greenwald et al., 2003).

3. Results of comparisons between all three conditions showed that when female participants were exposed to BS, they responded more quickly when women and competence shared the same response key (M=736ms; SD=105.83), compared to when they were not exposed to sexism (M=818ms; SD=118.46), $t(89)=8.96$, $p=.003$, $\eta p2=.048$. When exposed to HS (M=784ms; SD=101.41), their responses when women and competence shared a response key were not different compared to when they were not exposed to sexism, $t(89)=1.32$, $p=.252$, $\eta p2=.007$. Note that we expected HS to be similar to BS and differ from the no sexism condition. Nonetheless, HS (M=784ms; SD=101.41) did not differ from BS (M=736ms; SD=105.83), $t(89)=0.70$, $p=.792$, $\eta p2<.001$, and our analysis with weight contrasts showed that when both BS and HS were contrasted to the no sexism condition, participants showed weaker stereotype bias,

confirming our hypothesis. From a different analytical perspective, in these analyses also we found that women who were exposed to either BS or HS responded equally quickly when either male and competence or female and competence shared a response key, $t(89)=2.55$, p=.112, ηp2=.014 and $t(89)=1.80$, p=.182, ηp2=.010.

4. Note that the difference in *p* values between males and females is not an indication that the model fits the data better for females. Chi-square analyses are dependent on sample size and we had slightly unequal sample sizes (in this context sample size refers to the number of responses in the analysis: 13,393 for males and 16,189 for females). For this reason, we should rely on the *w* statistic which represents an effect size for chi-square analyses. This statistic controls for sample size and reveals identical effect sizes for males and females, suggesting that the model fits the data equally well for both sexes.

5. As we did with the follow-up GNAT response latency analyses reported in footnote 2, we also performed follow-up Quad model analyses with BS and HS separately for a more detailed description of our results. Female participants who were exposed to BS had stronger D (M=0.89) than female participants in the no sexism (M=0.80) condition, $\Delta\chi^2=58.03$, $p<.001$, $w=.07$, as well as female participants who were exposed to HS (M=0.84), $\Delta\chi^2=23.52$, $p<.001$, $w=.05$. Female participants who were exposed to HS also had stronger D than female participants in the no sexism condition, $\Delta\chi^2=6.65$, $p=.01$, $w=.03$. Female participants who were exposed to BS had lower G (M=0.39) than female participants in the no sexism (M=0.52) condition, $\Delta\chi^2=9.29$, $p=.002$, $w=.03$, as well as female participants who were exposed to HS (M=0.50), $\Delta\chi^2=6.26$, $p=.01$, $w=.02$. However, exposure to HS did not influence female participants' G relative to control, $\Delta\chi^2=0.20$, $p=.65$, $w<.01$. Exposure to BS marginally increased female participants' FAC (M=0.04) relative to both the no sexism (M=0.01) $\Delta\chi^2=3.26$, $p=.07$, $w=.02$, and HS

(*M*=0.01) conditions, $\Delta\chi^2$=3.33, *p*=.07, *w*=.02. However, exposure to HS did not influence female participants' FAC relative to no sexism condition, $\Delta\chi^2$<.01, *p*=.96, *w*<.01. Female participants' MAC and OB did not vary across BS, HS, and no sexism conditions, all $\Delta\chi^2$s<0.83, *p*s>.36, *w*s<.01. Male participants who were exposed to BS had stronger D (*M*=0.90) than male participants in the no sexism (*M*=0.86) condition, $\Delta\chi^2$=10.65, *p*=.001, *w*=.03, as well as male participants who were exposed to HS (*M*=0.85), $\Delta\chi^2$=11.88, *p*<.001, *w*=.04. However, exposure to HS did not influence male participants' D relative to the no sexism condition, $\Delta\chi^2$=0.46, *p*=.50, *w*<.01. Male participants who were exposed to BS had lower G (*M*=0.45) than male participants in the no sexism (*M*=0.57) condition, $\Delta\chi^2$=7.25, *p*=.007, *w*=.03, as well as male participants who were exposed to HS (*M*=0.57), $\Delta\chi^2$=5.81, *p*=.02, *w*=.03. However, exposure to HS did not influence male participants' G relative to the no sexism condition, $\Delta\chi^2$=0.20, *p*=.65, *w*<.01.Male participants' FAC, MAC, and OB did not vary across BS, HS, and no sexism conditions, all $\Delta\chi^2$s<2.74, *p*s>.10, *w*s<.02. Taken together, these follow-up analyses provide a more nuanced understanding of the process-level effects of the sexism manipulations than is reported in the main body of the paper. Overall, these analyses suggest that the BS manipulation had a larger effect than the HS manipulation; that the effect of the HS manipulation sometimes differed from the no sexism condition and sometimes did not; and that both sexism manipulations had larger effects on women than on men.

Table 1. Quad model estimated parameters for male and female participants.

	Male Participants		Female participants	
	No sexism	Sexism	No sexism	Sexism
FAC	0.00 [0.02]	0.00 [0.02]	0.01 [0.01]	0.02 [0.01]
MAC	0.02 [0.02]	0.03 [0.01]	0.02 [0.01]	0.02 [0.01]
D	0.86 [0.01]	0.88 [0.01]	0.80 [0.01]	0.87 [0.01]
OB	1.00 [2.08]	1.00 [0.95]	1.00 [1.34]	1.00 [0.40]
G	0.57 [0.04]	0.51 [0.04]	0.52 [.03]	0.45 [.02]

Notes: FAC = female-warmth associations, MAC = male-competence associations, D = Detection, G = Guessing, OB = Overcoming Bias. [Standard Error]

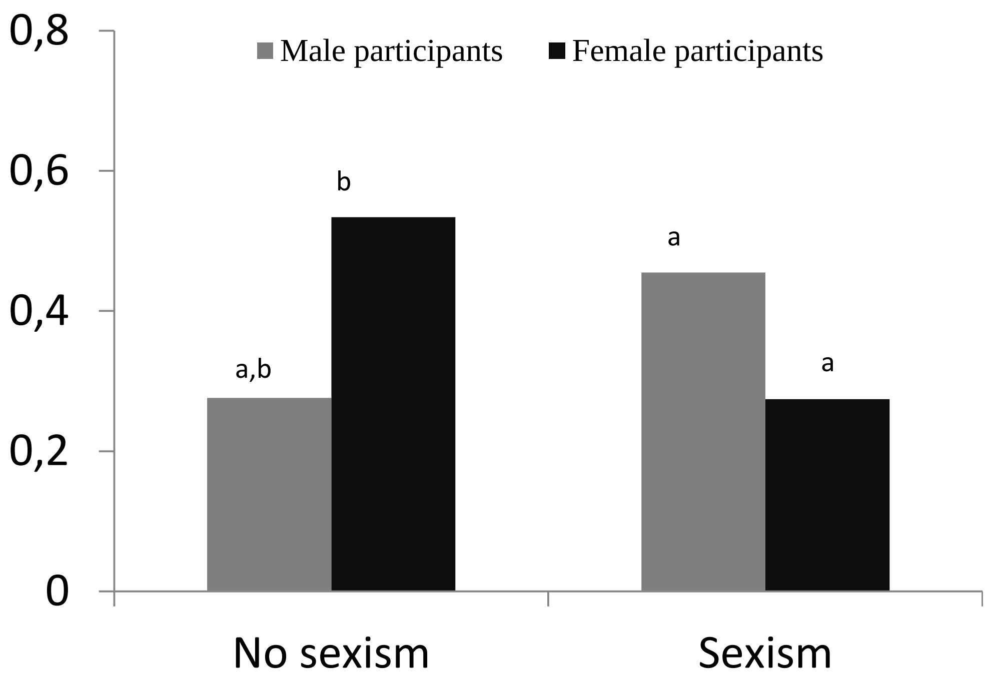

Figure 1. *D*-Scores (IAT) for male and female participants in no sexism and sexism exposure conditions (BS and HS collapsed). Higher scores reflect stronger gender stereotype bias. Bars with different subscripts differ significantly with $p < .050$. BS and HS response times were collapsed in the figure for ease of interpretation.

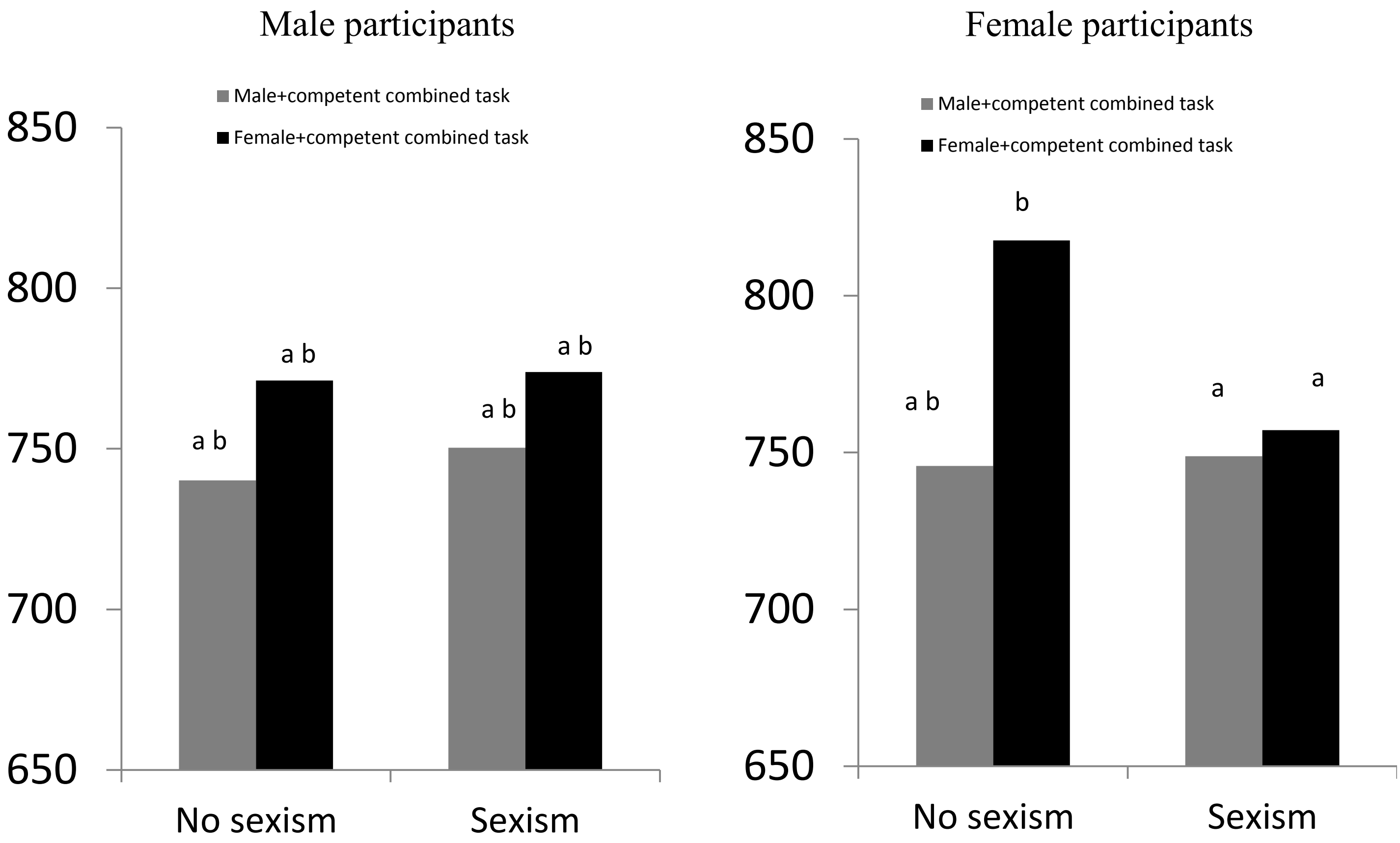

Figure 2. Response times on the GNAT for male and female participants in no sexism and sexism exposure conditions (BS and HS collapsed). Higher scores reflect slower responses to men and competence (grey bars) and women and competence (black bars) when they share a response key. Bars with different subscripts differ significantly with p < .050. BS and HS response times were collapsed in the figure for ease of interpretation.

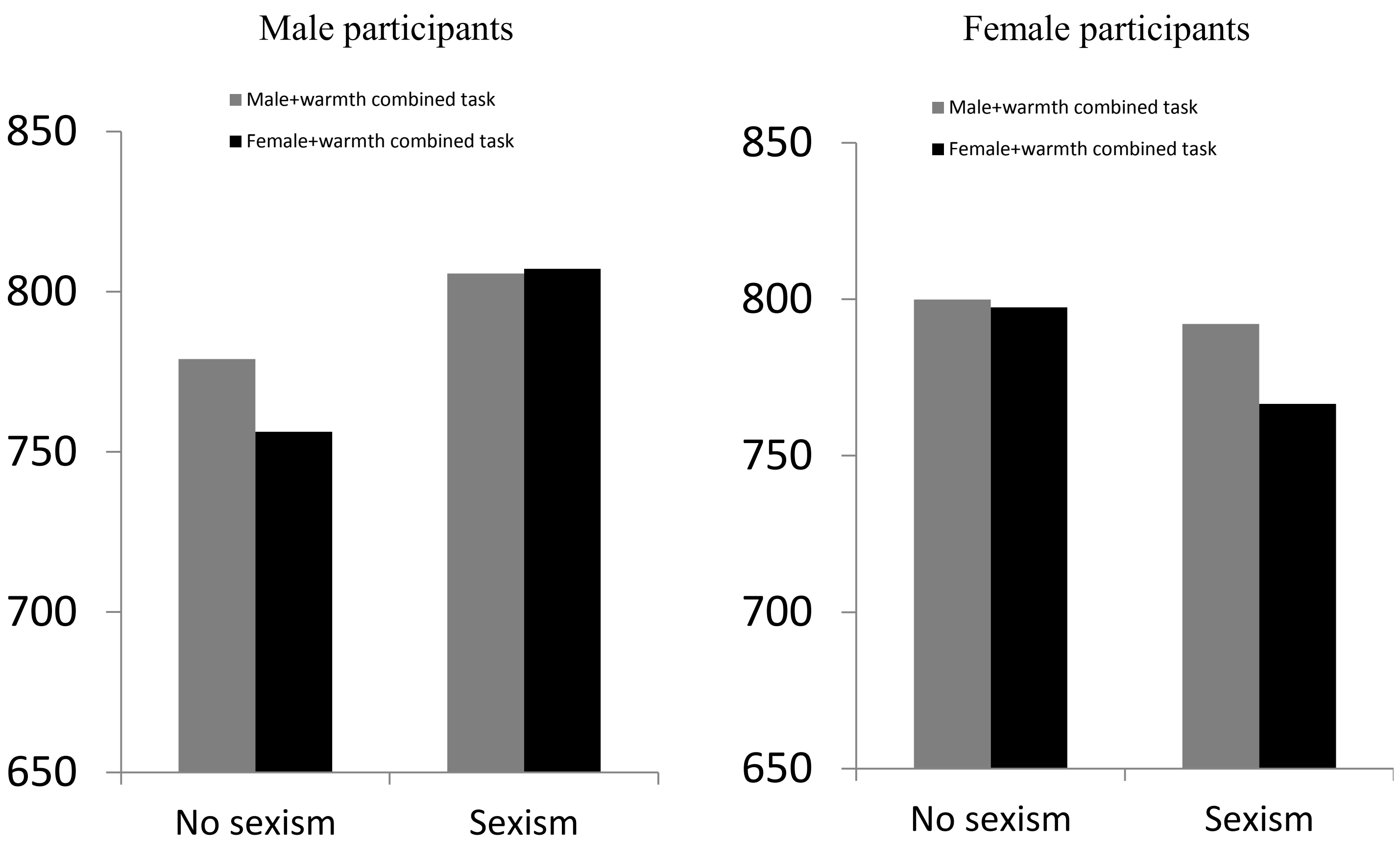

Figure 3. Response times on the GNAT for male and female participants in no sexism and sexism exposure conditions (BS and HS collapsed). Higher scores reflect slower responses to men and warmth (grey bars) and women and warmth (black bars) when they share a response key. BS and HS response times were collapsed in the figure for ease of interpretation.

Appendix

Benevolent sexism sentences (Experiments 1 and 2):

No matter how accomplished he is, a man is not truly complete as a person unless he has the love of a woman.

Many women have a quality of purity that few men possess

Women should be cherished and protected by men.

Every man ought to have a woman whom he adores.

A good woman should be set on a pedestal by her man.

Women, compared to men, tend to have a superior moral sensibility.

Hostile sexism sentences (Experiments 1 and 2):

Most women interpret innocent remarks or acts as being sexist

Women are too easily offended

Most women fail to appreciate fully all that men do for them.

Women seek to gain power by getting control over men

Women exaggerate problems they have at work

When women lose to men in a fair competition, they typically complain about being discriminated against.

No sexism sentences (Flowers; Experiment 1):

Water lilies are aquatic plants that can often cut the flow of water.

The wax plant bloom is formed by a star-shaped flower that looks as if it was made of wax or porcelain.

Tulips are thought to originate from Turkey and have a special meaning in the Netherlands.

Daisies may be used for tea such as chamomile.

Pagans worshiped the roses and believed in their mysticism according to the colour they displayed.

This flower's name comes from the mythological character Narcissus.

No sexism sentences (Keys; Experiment 2):

Despite not revealing to whom it belongs, each key always has a story to tell.

The key is an essential object in people's lives.

With so many models to choose from, keys can represent a personal taste.

Inspired in horror movies, the old style key is more easily associated with paranormal activity.

Due to their design, some keys can have a very exquisite look.

People need this object in their lives.