


Does Shared Gender Group Membership Mitigate the Effect of Implicit Bias Attributions on Accountability for Gender-Based Discrimination?

Personality and Social
Psychology Bulletin
2021, Vol. 47(9) 1343–1357
© 2020 by the Society for Personality
and Social Psychology, Inc
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0146167220965306
journals.sagepub.com/home/pspb


Natalie M. Daumeyer¹ , Ivuoma N. Onyeador¹ ,
and Jennifer A. Richeson¹

Abstract

Attributing gender discrimination to implicit bias has become increasingly common. However, research suggests that when discrimination is attributed to implicit rather than explicit bias, the perpetrators are held less accountable and deemed less worthy of punishment. The present work examines (a) whether this effect replicates in the domain of gender discrimination, and (b) whether sharing a group membership with the victim moderates the effect. Four studies revealed that both men and women hold perpetrators of gender discrimination less accountable if their behavior is attributed to implicit rather than explicit bias. Moreover, women held male (Studies 1–3), but not female (Study 4), perpetrators of gender discrimination more accountable than did men. Together, these findings suggest that while shared gender group membership may inform judgments of accountability for gender discrimination, it does not weaken the tendency to hold perpetrators less accountable for discrimination attributed to implicit, compared with explicit, bias.

Keywords

bias attribution, gender dynamics, social identity, implicit versus explicit bias, discrimination

Received August 27, 2019; revision accepted September 18, 2020

Gender discrimination remains a pervasive issue in American society. Women earn less money than men for the same work (PayScale, 2019) and 42% of women report experiencing gender discrimination in the workplace (Parker & Funk, 2017), a problem that is even more prevalent in male-dominated industries, such as technology (Ziv, 2016). Furthermore, it is becoming increasingly popular for people to attribute instances of gender discrimination to the implicit bias of its perpetrators (Altmire, 2019; Manne, 2019; Mundy, 2017). Although it is important for people to understand the subtle ways that biases can affect our behaviors, attributions of discrimination to implicit bias can have unforeseen consequences. For instance, a growing body of research suggests that when discriminatory behaviors are attributed to implicit or unconscious bias, rather than to explicit or conscious bias, people hold the perpetrators of those acts less accountable (Cameron et al., 2010; Daumeyer et al., 2019; Redford & Ratliff, 2016). Specifically, Daumeyer et al. (2019) exposed participants to communications of scientific findings, revealing the discriminatory effects of implicit, or rather, explicit bias. Across four studies, participants held perpetrators less accountable for discriminatory behavior when it was

attributed to their implicit, rather than explicit, attitudes (meta-analytic $d = .41$, $z = 7.08$, $p < .001$). This reduced accountability for implicit bias effect was observed in two different contexts (medical and police interactions), across three different biases (political, age-based, and racial), and, somewhat surprisingly, was not attenuated when the consequences of the discrimination were especially severe (i.e., premature death).

In other words, the tendency to reduce judgments of accountability for discriminatory acts that have been attributed to implicit rather than explicit bias appears to be quite robust. Given rising concerns about gender inequality (e.g., the Times Up and #MeToo movements), however, it is unclear whether people, in general, and women, in particular, will be willing to hold (male) perpetrators any less accountable for gender-based discrimination that is attributed to

¹Yale University, New Haven, CT, USA

Corresponding Author:

Natalie Daumeyer, Department of Psychology, Yale University, 2 Hillhouse Ave, New Haven, CT 06520-8205, USA.
Email: natalie.daumeyer@yale.edu

implicit bias. Because women share a relevant group membership with the most likely victims of gender discrimination, they may be reluctant to reduce their judgments of accountability for male perpetrators based on implicit bias attributions. The present work sought to examine this possibility. Specifically, we test whether (a) perpetrators are held less accountable for gender discrimination attributed to implicit bias compared with explicit bias, and (b) this tendency is observed among both male and female perceivers.

Shared Gender Group Membership and Perceptions of Discrimination

When evaluating instances of wrongdoing, people consider both the intent of the perpetrator as well as the harm to the victim (Malle et al., 2014; McNamara et al., 2019; Nadler & McDonnell, 2011). Perpetrators are held more accountable and punished more when they are thought to be acting intentionally and the more harmful the outcome (Cushman, 2008; Ginther et al., 2016; Monroe & Malle, 2018). Because implicit bias is understood as being largely unconscious and/or giving rise to behavior that is unintentional, people who engage in discrimination attributed to implicit bias are held less culpable than those whose discriminatory behavior is attributed to explicit bias (Cameron et al., 2010; Daumeyer et al., 2019; Onyeador, 2017; Redford & Ratliff, 2016). Because our group memberships affect how we perceive and evaluate the behavior of other people (Brewer & Kramer, 1985; Stürmer et al., 2005; Tanis & Postmes, 2005; Weisbuch & Ambady, 2008; Xu et al., 2009), however, sharing a relevant group membership with either the perpetrator or the victim of acts of discrimination might change these patterns of accountability.

Members of low-status groups, including women in many societal domains, are often more sensitive to acts of discrimination faced by fellow ingroup members, compared with members of high-status groups. For example, women perceive gender discrimination as more prevalent and more harmful than do men (Guttek et al., 1996; Swim et al., 2003). Consequently, when considering cases of gender discrimination attributed to implicit bias, sharing a group membership with the victim may lead female perceivers to focus less on the intentions of the perpetrator, compared with male perceivers and, perhaps, more on the harm incurred by the victim (Simon, Moss, & O'Brien, 2019). Focusing less on the intentions of the perpetrator should, in turn, result in greater assessments of perpetrator accountability, even when gender discrimination is attributed to implicit bias. Consequently, women may not hold male perpetrators any less accountable for discriminatory behavior toward other women when it is attributed to implicit, rather than explicit, gender-stereotypical biases and beliefs.

In addition to these relatively cognitive explanations for why shared gender group membership may result in a differential tendency to reduce perpetrator accountability for discrimination born of implicit bias, there are also motivational

explanations. Specifically, seeing or hearing about an ingroup member experiencing or perpetuating discrimination is threatening to one's social identity (Branscombe et al., 1999; Major et al., 2002; Major & O'Brien, 2005). That is, sharing a group membership with either the perpetrator or victim of discrimination may be experienced as a threat to the value of one's social identity (Branscombe et al., 1999). For instance, men may experience a threat to the perceived morality of their group upon learning that a man has discriminated against a woman (Branscombe et al., 1999). One way to diffuse such a threat is to downplay the immorality of the behavior (e.g., to justify or explain the behavior). Thus, men may look for mitigating information that could account for the gender discrimination of other men (Jones & Davis, 1965); implicit bias may offer just such a compelling explanation. In other words, perceivers who share a relevant group membership with the perpetrator of discrimination may be especially eager to rely on an implicit bias attribution as an excuse to hold the perpetrator less accountable.

In cases of gender discrimination by male perpetrators toward female targets, women may be motivated to do just the opposite. Because sharing a group membership with the victim of the discrimination is a threat to their group's value and/or esteem (Branscombe et al., 1999; Major & O'Brien, 2005), women may be eager to hold the perpetrator accountable (Malle et al., 2014; Monroe & Malle, 2018), if only to reduce the likelihood that the discriminatory behavior will persist. In other words, sharing a gender group membership with the victim of discrimination may increase people's motivation to hold (outgroup) perpetrators accountable and, thus, result in less discounting of perpetrator accountability for acts of discrimination that are attributed to implicit, rather than explicit, bias.

The Present Work

The primary aims of the present work are to explore (a) whether perpetrators of gender discrimination are held less accountable when their behavior is attributed to implicit compared with explicit bias, and (b) whether the magnitude of this reduced accountability for implicit bias effect is moderated by shared gender group membership with the perpetrator or victim of the discrimination. To address these aims, we conducted four studies examining the extent to which male and female perceivers hold male (Studies 1–3) and female (Study 4) perpetrators less accountable for gender discrimination when it is attributed to their implicit, compared with their explicit, bias. Replicating past work in other domains of discrimination (Cameron et al., 2010; Daumeyer et al., 2019; Redford & Ratliff, 2016), we expected participants to hold perpetrators less accountable and be less supportive of punishing them for acts of gender discrimination when attributed to their implicit, rather than explicit, bias. We also expected to observe a main effect of participant gender. Specifically, in Studies 1 to 3, we expected women, who share a group

membership with the victim, to hold the male perpetrator more accountable and be more supportive of punishing him than men, who share a group membership with the perpetrator. In Study 4, we expected men, who in this case share a group membership with the victim, to hold the female perpetrator more accountable and support punishing her more than women, who share a group membership with the perpetrator. In addition, we expected to observe a significant interaction between the bias attribution condition and participant gender (i.e., shared gender group membership). Specifically, we expected the magnitude of the reduced accountability for discrimination attributed to implicit, versus explicit, bias effect (i.e., the main effect of bias attribution) to be smaller among participants who share a group membership with the victim, compared with participants who share a group membership with the perpetrator, of the discrimination.

Study 1

Design and Materials

This study had a 2 (Bias Attribution: Explicit vs. Implicit) \times 2 (Participant Gender: Men vs. Women) between-subjects design. All materials and data for the four studies presented here can be found at https://osf.io/tybvq/?view_only=137c871f85f3468eafc263d1324708b0. Participants responded to all dependent measures on scales from 1 (*strongly disagree*) to 7 (*strongly agree*).

Participants. Through TurkPrime (Litman et al., 2017), 300 people in the United States completed the study in exchange for US\$1.00. Three people reported that they did not answer carefully or accurately, three reported being suspicious of the article manipulation, and two reported their gender as “other” leaving a total of 292 participants for analysis (51.4% women, 79.1% White, $M_{\text{age}} = 36.38$, 145 in the explicit condition). This sample gave us greater than 80% power to detect an effect of $\eta^2_{\text{partial}} = .026$, which is similar to the effect size estimates observed in related work (Daumeyer et al., 2019).

Bias manipulation. Participants read an ostensible news article detailing research about gender discrimination against women working in science, technology, engineering, and mathematics (STEM) companies (see Ziv, 2016). The discriminatory behaviors were attributed to either the explicit or implicit biases and beliefs held by the male employees and managers at the companies. In the implicit condition, participants were given a brief definition of implicit bias: “Implicit biases are attitudes or stereotypes that affect our actions and decisions in ways that we are typically not even aware.” The bias in the article used in this condition was always referred to as “implicit.”

Accountability. Perceptions of perpetrator accountability were measured with a six-item scale ($\alpha = .81$, for example,

“People should be held responsible for any gender biases they have that may impact their coworkers.”).

Punishment. Support for punishing perpetrators was measured using a two-item scale ($\alpha = .69$, for example, “Managers who consistently demonstrate gender biases should be penalized or even demoted.”).

Other measures. In addition to accountability and punishment, as in Daumeyer et al. (2019), we measured participants’ level of concern about the bias they read about and support for reform efforts to mitigate the bias, as well as several potential moderator variables (e.g., bias perception; Uhlmann & Nosek, 2012, motivation to respond without prejudice; Plant & Devine, 1998). Because they were not measured across all studies, analyses of concern and reform are provided in the Supplemental Material available online. None of the measured individual difference variables consistently moderated the main effects on any of the outcome variables and, thus, they are not discussed further. Study 1 did not have a bias attribution manipulation check.

Procedure

After providing informed consent, participants completed relevant demographic information (e.g., age, gender, race, and political ideology). Participants were subsequently randomly assigned to a bias attribution condition, within gender. After reading the article for their condition, participants completed the primary outcome measures—accountability and punishment, followed by the other measures, potential moderator variables, attention checks, and additional demographic information (e.g., whether they currently work for a STEM company).¹ Finally, participants reported whether they responded carefully and accurately, and then they were thanked, debriefed, and paid.

Results

To test our predictions, a 2 (Bias Attribution: Explicit vs. Implicit) \times 2 (Participant Gender: Men vs. Women) analysis of variance (ANOVA) was conducted on the dependent variables. Mean values and standard deviations for the dependent variables are provided in Table 1.

Accountability. Consistent with predictions and past work (Daumeyer et al., 2019), participants in the implicit bias condition held the perpetrators less accountable than participants in the explicit bias condition, $F(1, 288) = 4.80$, $p = .029$, $\eta^2_{\text{partial}} = .016$. Also consistent with predictions, analyses revealed a main effect of participant gender, $F(1, 288) = 17.99$, $p < .001$, $\eta^2_{\text{partial}} = .059$. Women held the male perpetrators more accountable for their discrimination than did men. Contrary to predictions, participant gender did not moderate the effect of bias attribution on perceived accountability, $F(1, 288) = 0.14$, $p = .71$, $\eta^2_{\text{partial}} < .001$.

Table 1. Mean Values and Standard Deviations by Gender and Condition for Studies 1 and 2.

	Study 1									Study 2								
	All			Men			Women			All			Men			Women		
	M	SD	N	M	SD	N	M	SD	N	M	SD	N	M	SD	N	M	SD	N
<i>Accountability</i>																		
Explicit	5.36 ^a	1.04	145	5.12	1.15	71	5.58	0.88	74	5.78 ^a	0.91	541	5.68	0.93	280	5.89	0.87	261
Implicit	5.10 ^b	1.07	147	4.81	1.14	71	5.36	0.93	76	5.53 ^b	0.92	539	5.35	0.98	274	5.72	0.80	265
Total	5.22	1.06	292	4.96 _a	1.15	142	5.47 _b	0.91	150	5.66	0.92	1,080	5.52 _a	0.97	554	5.80 _b	0.84	526
<i>Punishment</i>																		
Explicit	5.28	1.35	145	4.97	1.41	71	5.57	1.22	74	5.51 ^a	1.01	541	5.33	1.08	280	5.71	0.90	261
Implicit	5.01	1.43	147	4.70	1.53	71	5.30	1.28	76	5.17 ^b	1.10	539	5.02	1.21	274	5.33	0.94	265
Total	5.14	1.40	292	4.84 _a	1.48	142	5.43 _b	1.25	150	5.34	1.07	1,080	5.18 _a	1.15	554	5.52 _b	0.94	526

Note. Condition mean values that differ significantly at the $p < .05$ level are designated with different superscripts in the All column. Participant gender effects that differ significantly at the $p < .05$ level are designated with different subscripts in the Total row for men and women.

Punishment. The results for punishment mirrored those for accountability. There was a nonsignificant trend for participants in the implicit bias condition to support punishment less than participants in the explicit bias condition, $F(1, 288) = 2.85, p = .092, \eta^2_{\text{partial}} = .010$. The main effect of participant gender was significant, $F(1, 288) = 13.83, p < .001, \eta^2_{\text{partial}} = .046$. Women were more supportive of punishing male perpetrators of gender discrimination than were men. And, again, there was no interaction between bias attribution and participant gender, $F(1, 288) = 0.00, p = .99, \eta^2_{\text{partial}} < .001$.

Discussion

Study 1 revealed that both men and women reduce the extent to which they hold male perpetrators of gender discrimination accountable for their behavior when it is attributed to their implicit, rather than explicit, attitudes and beliefs. That said, women did tend to hold the male perpetrators more accountable overall, compared with their male counterparts. The results for punishment revealed this same general pattern. Together, these findings suggest that sharing a gender with the victim of gender discrimination does not *necessarily* reduce the extent to which perceivers hold perpetrators less accountable for discrimination attributed to implicit, compared with explicit, bias. Prior to speculating why this pattern of results may have emerged, we thought it important to reexamine the potential for moderation due to sharing a gender group membership with the victim (rather than perpetrator) of gender discrimination in a study that is both sufficiently powered to detect the small effects observed here.

Study 2

Study 2 sought to examine the robustness of the effects in Study 1 in a sample large enough to observe an interaction between the bias attribution condition and participant gender.

The present study also examined the generalizability of the effects of Study 1 to an interpersonal scenario in which an actor behaves in a discriminatory way toward a single victim, thus, increasing the potential for participants to identify with either the victim or the perpetrator in the situation.

Design and Materials

As in Study 1, this study had a 2 (Bias Attribution: Explicit vs. Implicit) \times 2 (Participant Gender: Men vs. Women) between-subjects design. Study 2 was pre-registered at aspredicted.org (<http://aspredicted.org/blind.php?x=vd7fe7>).

Participants. Using TurkPrime (Litman et al., 2017), 1,100 people in the United States completed the study in exchange for US\$0.80. Thirteen people reported that they did not answer carefully or accurately and seven reported a gender other than “man” or “woman,” leaving a total of 1,080 participants for analysis (48.7% women, 68.1% White, $M_{\text{age}} = 34.43$, 541 in the explicit condition). This final sample size had more than 95% power to detect an interaction even assuming a small effect size ($\eta^2_{\text{partial}} = .01$).

Bias manipulation. Participants read a scenario about a woman named Naomi who was experiencing gender discrimination at her tech company. We manipulated the attribution for the discrimination by telling participants that an outside consulting firm reviewed the case and, in their report, they either stated,

Naomi’s experience of discrimination seems to be caused by her manager’s **conscious** negative stereotypes about the technical ability of women compared to men. In other words, Naomi’s manager is aware that he holds negative stereotypes about gender and technical ability

in the explicit condition, or

Naomi's experience of discrimination seems to be caused by her manager's **unconscious** negative stereotypes about the technical ability of women compared to men. In other words, Naomi's manager is not aware that he holds negative stereotypes about gender and technical ability

in the implicit condition. For the full scenarios see https://osf.io/tybvq/?view_only=137c871f85f3468eafc263d1324708b0.

Accountability. Perceptions of accountability ($\alpha = .77$) were measured similar to that in Study 1, but items were adjusted to be about the manager in the scenario rather than all managers (e.g., "The manager should be held responsible for any gender bias he has that may impact Naomi.").

Punishment. Support for punishing the manager was measured similar to that in Study 1, but with five items ($\alpha = .81$, for example, "Naomi's manager should be put on probation.").

Manipulation check. Participants completed a one-item manipulation check asking them to complete the sentence, "In the beginning of the study, the instance of gender discrimination in the report was attributed to . . ." with the following options: the manager's *unconscious* negative stereotypes about women's technical abilities, the manager's *conscious* negative stereotypes about women's technical abilities, a lack of female managers, I'm not sure, and the instance was not attributed to anything. In all, 193 people answered the manipulation check incorrectly: 84 people (15.5%) in the explicit bias condition and 109 people (20.2%) in the implicit bias condition. The number of people who failed the manipulation check was significantly higher in the implicit bias condition, $\chi^2(1, N = 1,080) = 4.06, p = .044$. The results presented here include participants who answered the manipulation check incorrectly; however, all main effects and interactions on the outcome variables are the same when these participants are excluded.

Other measures. In addition to our primary outcomes, we again measured support for reform as well as the Ambivalent Sexism Inventory (ASI; Glick & Fiske, 1996) to explore potential moderation by hostile and benevolent sexism. These analyses are provided in the Supplemental Material available online. The results reported below are robust to the inclusion of benevolent and hostile sexism as covariates.

Procedure

After providing informed consent, participants read the scenario about Naomi experiencing gender discrimination at work. They subsequently completed the primary outcome measures: perceived accountability and support for punishment, followed by the support for reform efforts measure, the bias attribution manipulation check, and the ASI. Finally, they reported on their demographic information, indicated

whether they responded carefully and accurately, and were then thanked, debriefed, and paid.

Results

A 2 (Bias Attribution: Explicit vs. Implicit) \times 2 (Participant Gender: Men vs. Women) ANOVA was conducted on the primary dependent variables. Mean values and standard deviations are provided in Table 1.

Accountability. Replicating Study 1, the effect of bias condition was significant, $F(1, 1076) = 19.77, p < .001, \eta^2_{\text{partial}} = .018$. Consistent with predictions, participants in the implicit condition held the perpetrator less accountable than participants in the explicit condition. The expected main effect of gender also emerged, $F(1, 1076) = 28.03, p < .001, \eta^2_{\text{partial}} = .025$. Women held the perpetrator more accountable than men. Replicating Study 1, the interaction between participant gender and bias attribution was not significant, $F(1, 1076) = 2.24, p = .14, \eta^2_{\text{partial}} = .002$.

Punishment. Similar to the accountability results, both the bias attribution main effect, $F(1, 076) = 29.69, p < .001, \eta^2_{\text{partial}} = .027$, and the main effect of participant gender, $F(1, 076) = 29.15, p < .001, \eta^2_{\text{partial}} = .026$, emerged significant. Participants in the implicit bias condition thought that the perpetrator should be punished less than did participants in the explicit bias condition and women were more supportive of punishment than were men. As before, the interaction between participant gender and bias attribution condition was not significant, $F(1, 1076) = 0.26, p = .61, \eta^2_{\text{partial}} < .001$.

Discussion

Replicating Study 1, as well as past work (Daumeyer et al., 2019), participants held a perpetrator less accountable, and supported punishing him less, for acts of discrimination that were attributed to his implicit, rather than explicit, attitudes. Furthermore, women held the male perpetrator more accountable and supported punishing him more than men. Somewhat surprisingly, however, participant gender did not significantly moderate the tendency to hold the perpetrator less accountable for discrimination attributed to his implicit, rather than explicit, attitudes.

Recall that we thought that participant gender might predict a differential tendency to hold the perpetrator accountable for gender discrimination attributed to implicit bias, primarily due to cognitive and motivational processes stemming from sharing a relevant group membership with the victim, rather than the perpetrator, of the discrimination. Specifically, sharing a gender with the victim of the discrimination may shift female perceivers' focus from the intentions of the male perpetrator and, perhaps, toward the experiences of the female victim, ultimately disrupting the processes thought to give rise to the reduced accountability for implicit

bias effect. Although the lack of moderation of this effect by participant gender observed in the present study suggests that our reasoning regarding the potential effects of shared gender group membership may not be correct, it is certainly possible that the main effects of participant gender found could be due to a differential tendency to take on the perspective of the victim and/or the perpetrator in the discrimination scenario. Study 3 aimed to examine this possibility.

Study 3

The aim of Study 3 was to replicate and extend Study 2 by adding measures of perspective-taking and sympathy for the victim and perpetrator. This allows us to test whether the previously observed main effect of gender on accountability and punishment is due, at least in part, to the tendency for women to take the perspective of, and sympathize with, the female victim more than male participants.

Design and Materials

Study 3 had a 2 (Bias Attribution: Explicit vs. Implicit) \times 2 (Participant Gender: Men vs. Women) between-subjects design. Study 3 was pre-registered at [aspredicted.org](http://aspredicted.org/blind.php?x=by9nz7) (<http://aspredicted.org/blind.php?x=by9nz7>).

Participants. Using TurkPrime (Litman et al., 2017), 656 people in the United States completed the study in exchange for US\$0.75. The sample size was determined by using a power analysis on the smallest significant effect size from Study 2. Eleven participants were excluded from analyses: four for admitting they did not answer carefully or accurately and seven who reported a gender other than “man” or “woman.” Therefore, we had a total of 645 participants for analysis (55.7% women, 71.8% White, $M_{\text{age}} = 36.47$, 320 in explicit condition). This final sample provided more than 95% power to detect an interaction even assuming a small effect size ($\eta^2_{\text{partial}} = .02$).

Discrimination scenario and bias attribution manipulation. The scenario described in Study 2 was used in the present study and, thus, the bias attribution manipulation was also as described previously.

Primary outcome variables. Perceptions of accountability and support for punishment were measured as described in Study 2.

Manipulation check. The manipulation check was measured the same as in Study 2. In Study 3, 95 people (14.7%) failed the manipulation check: 43 people (13.4%) in the explicit condition and 52 people (16.0%) in the implicit condition. The number of people who failed the manipulation check did not differ by condition, $\chi^2(1, N = 645) = 0.84, p = .359$. All main effects and interactions on the outcome variables

are the same when these participants are excluded from analyses.

Perspective-taking. We included measures of perspective-taking with the victim (Naomi) and the perpetrator (the male manager) to test whether perspective-taking differed as a function of participant gender and, perhaps, predicted differential judgments of perpetrator accountability and support for punishment. Specifically, participants responded to two face-valid statements: “When I read the scenario, I imagined myself in the manager’s position” and “When I read the scenario, I imagined myself in Naomi’s position,” each on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*). Items were presented in random order across participants.

Sympathy. We also measured the extent to which participants felt sympathy toward the victim and the perpetrator. These two measures were determined through a factor analysis using Varimax rotation (see Supplemental Material available online). Sympathy toward the victim ($\alpha = .90$) measured the extent to which participants felt sympathy, compassion, and concern toward Naomi and anger, outrage, and irritation toward the manager. Sympathy toward the perpetrator ($\alpha = .87$) measured the extent to which participants felt sympathy, compassion, and concern toward the manager and anger, outrage, and irritation toward Naomi. These emotions were rated on scales from 1 (*not at all*) to 7 (*very much*). The order of the scales and items within the scales were randomized across participants.

Other measures. We also measured participants’ level of support for efforts to redress the harm incurred by the victim with three items ($\alpha = .58$, for example, “Naomi should be financially compensated by the company.”). Analyses for support for redress are provided in the Supplemental Material available online.

Procedure

After providing their informed consent, participants read the discrimination scenario and then completed the measures of perspective-taking, sympathy, accountability, punishment, and redress, in that order. They next completed the same bias attribution condition manipulation check item from Study 2, after which they reported on their demographic information and whether they answered carefully and accurately, prior to being thanked, debriefed, and paid.

Results

Each dependent variable was submitted to a 2 (Bias Attribution: Explicit vs. Implicit) \times 2 (Participant Gender: Male vs. Female) ANOVA. Mean values and standard deviations for accountability and punishment are provided in Table 2.

Table 2. Mean Values and Standard Deviations by Gender and Condition for Studies 3 and 4.

	Study 3									Study 4								
	All			Men			Women			All			Men			Women		
	M	SD	N	M	SD	N	M	SD	N	M	SD	N	M	SD	N	M	SD	N
<i>Accountability</i>																		
Explicit	5.82 ^a	0.99	320	5.58	1.03	143	6.02	0.91	177	5.31 ^a	0.99	344	5.33	1.02	162	5.28	0.96	182
Implicit	5.65 ^b	1.00	325	5.45	1.12	143	5.81	0.87	182	4.78 ^b	0.98	346	4.82	1.00	120	4.76	0.97	226
Total	5.74	1.00	645	5.52 _a	1.08	286	5.91 _b	0.89	359	5.04	1.02	690	5.11	1.04	282	4.99	1.00	408
<i>Punishment</i>																		
Explicit	5.55 ^a	1.09	320	5.37	1.15	143	5.70	1.02	177	4.44 ^a	1.18	344	4.53	1.16	162	4.37	1.19	182
Implicit	5.33 ^b	1.13	325	5.13	1.26	143	5.49	0.98	182	3.78 ^b	1.13	346	3.89	1.18	120	3.73	1.10	226
Total	5.44	1.11	645	5.25 _a	1.21	286	5.59 _b	1.00	359	4.11	1.20	690	4.26	1.21	282	4.01	1.18	408

Note. Condition mean values that differ significantly at the $p < .05$ level are designated with different superscripts in the All column. Participant gender effects that differ significantly at the $p < .05$ level are designated with different subscripts in the Total row for men and women.

Accountability. Replicating the previous studies, analyses revealed that participants in the implicit bias condition held the perpetrator less accountable than participants in the explicit bias condition, $F(1, 641) = 4.86, p = .028, \eta^2_{\text{partial}} = .008$. And, as in the previous two studies, women held the (male) perpetrator more accountable than men, $F(1, 641) = 26.72, p < .001, \eta^2_{\text{partial}} = .040$. Participant gender again did not moderate the effect of bias attribution on accountability, $F(1, 641) = 0.32, p = .57, \eta^2_{\text{partial}} < .001$.

Punishment. Similarly, analyses revealed that participants in the implicit bias condition were less supportive of punishing the perpetrator than those in the explicit bias condition, $F(1, 641) = 6.64, p = .010, \eta^2_{\text{partial}} = .010$. In addition, women were more supportive of punishing the male perpetrator than were men, $F(1, 641) = 15.32, p < .001, \eta^2_{\text{partial}} = .023$. Again, the interaction between participant gender and bias attribution condition was not significant, $F(1, 641) = 0.03, p = .88, \eta^2_{\text{partial}} < .001$.

Perspective-taking. Analyses revealed that participant gender significantly predicted perspective-taking with both the victim, $F(1, 641) = 30.65, p < .001, \eta^2_{\text{partial}} = .046$, and perpetrator, $F(1, 641) = 34.53, p < .001, \eta^2_{\text{partial}} = .051$. Women ($M = 5.81, SD = 1.29$) took the female victim’s perspective more than men ($M = 5.20, SD = 1.48$), whereas men ($M = 3.62, SD = 1.78$) took the male perpetrator’s perspective more than women ($M = 2.80, SD = 1.79$). In addition, there was a significant effect of bias attribution on perspective-taking with the perpetrator, $F(1, 641) = 5.70, p = .017, \eta^2_{\text{partial}} = .009$. Participants reported greater levels of perspective-taking with the perpetrator when the discrimination was attributed to implicit ($M = 3.34, SD = 1.86$) rather than explicit ($M = 2.98, SD = 1.79$) bias. No other effects were reliable ($ps > .09$).

Sympathy. A similar pattern emerged for participant gender on sympathy with the victim and perpetrator. Women

($M = 5.65, SD = 1.20$) expressed more sympathy for the (female) victim than did men ($M = 5.06, SD = 1.38$), $F(1, 641) = 33.90, p < .001, \eta^2_{\text{partial}} = .050$. Similarly, men ($M = 2.60, SD = 1.42$) expressed more sympathy for the (male) perpetrator than did women ($M = 2.02, SD = 1.22$), $F(1, 641) = 31.15, p < .001, \eta^2_{\text{partial}} = .046$. No other effects were reliable ($ps > .20$).

Mediation analyses. We next sought to test whether the more participants reported taking the victim’s perspective and feeling sympathy toward her (and the less they took the perpetrator’s perspective and felt sympathy for him), the more they held the perpetrator accountable and supported punishing him. Specifically, we used PROCESS (5,000 bootstrapped resamples, Model 4; Hayes, 2013) to test whether victim perspective-taking, perpetrator perspective-taking, victim sympathy, and/or perpetrator sympathy mediated the relationship between participant gender (men = -0.5 , women = 0.5) and the outcome variables, with bias attribution condition included as a covariate. As depicted in Figure 1, women’s tendency to take the victim’s perspective more, $B = 0.60, SE = 0.11, p < .001$, feel more sympathy toward the victim, $B = 0.59, SE = 0.10, p < .001$, and feel less sympathy for the perpetrator, $B = -0.58, SE = 0.10, p < .001$, all compared with men’s responses on these measures and statistically explained their tendency to hold the male perpetrator more accountable compared with men (total indirect effect: $B = 0.46, SE = 0.06, 95\% \text{ confidence interval [CI]} = [.34, .59]$). And, the direct effect of participant gender on accountability was no longer significant with the mediators in the model, $B = -0.06, SE = 0.05, p = .25$ (total effect: $B = 0.40, SE = 0.08, p < .001$).

A similar pattern emerged for punishment (total indirect effect: $B = 0.41, SE = 0.07, 95\% \text{ CI} = [.27, .54]$). Relative to men, women expressed greater sympathy for the victim, $B = 0.59, SE = 0.10, p < .001$, and lower sympathy for the perpetrator, $B = -0.58, SE = 0.10, p < .001$, which significantly predicted and statistically mediated the effect of

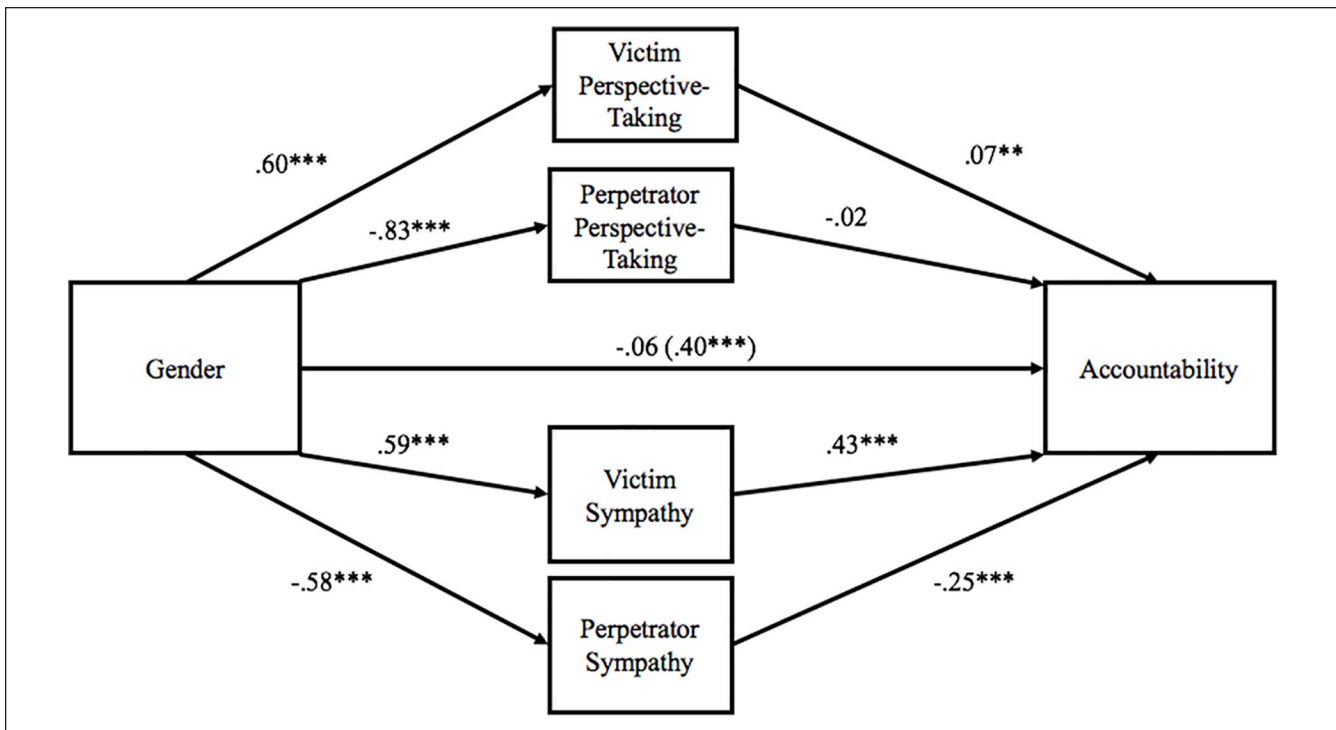


Figure 1. Mediation model depicting the effect of gender on accountability in Study 3.

Note. Bias attribution condition was included as a covariate.

** $p < .01$. *** $p < .001$.

participant gender on support for punishing the perpetrator ($B = 0.54$, $SE = 0.03$, $p < .001$, and $B = -0.11$, $SE = 0.03$, $p < .001$, for victim and perpetrator sympathy, respectively). And, once again, the direct effect of participant gender was no longer significant, $B = -0.07$, $SE = 0.07$, $p = .33$ (total effect: $B = 0.34$, $SE = 0.09$, $p < .001$).

Discussion

Study 3 replicated the pattern of effects found in Studies 1 and 2; when sexist behaviors were attributed to the male perpetrator's implicit bias, he was held less accountable and perceived to be less worthy of punishment than when those same behaviors were attributed to his explicit bias. Both men and women revealed this reduced accountability for implicit bias effect. In addition, women, perhaps because they share a group membership with the victim of the discrimination described in the scenario, held the perpetrator more accountable and were more supportive of punishing him than were men, perhaps because they share a group membership with the perpetrator. Consistent with this explanation, mediation analyses suggested that perspective-taking and sympathy help explain the relationship between participant gender and the outcome variables: accountability and punishment. Specifically, the present study offered initial evidence that women held the male perpetrator more accountable than men, in part because women took the perspective of and

sympathized with the female victim more than men did as well as because women sympathized with the male perpetrator less than men did. Similarly, women tended to be more supportive of punishing the perpetrator than men because they sympathized more with the victim and less with the perpetrator than men did.

Although shared group membership is one plausible explanation for this pattern, it is not the only one. It is also possible that the differential patterns of perspective-taking (and felt sympathy) are due to processes related to gender (and gender socialization) that are independent of sharing a relevant group membership with the perpetrator or victim of the discrimination. Notably, it is possible that women were more likely to take the perspective of the victim of the discrimination and feel more sympathy toward her simply because of gender socialization processes that encourage women, more than men, to express sympathy toward others in the wake of harm (Hess et al., 2000; Lithari et al., 2010). Alternatively, these gendered patterns of differential perspective-taking and sympathy could reflect the effects of membership in a group that is a common and traditional target of societal discrimination (Guttek et al., 1996). That is, it could be the case that, regardless of whether they share a group membership with the victims (or perpetrators) of specific instances of discrimination, women will be more likely to adopt the perspective of the victim, compared with men, and, in turn, hold the perpetrator more accountable.

We cannot disambiguate these possibilities in the present study, of course, because participant gender is conflated with shared group membership with the victim of the discrimination (and, perhaps importantly, not sharing one with the perpetrator). In Study 4, however, we switch the contingency between participant gender and perpetrator versus victim gender. Specifically, participants evaluate a case of gender discrimination by a female perpetrator toward a male victim. This arrangement allows us to explore whether women continue to adopt the perspective of the male victim of discrimination, sympathize with him more, and hold the female perpetrator more accountable, compared with men, consistent with both the gender socialization and “traditional target of discrimination” explanations. Or, rather, consistent with the shared group membership explanation, women may take the perspective of the female perpetrator more, sympathize with her more and, thus, hold her less accountable compared with men, who, in turn, take the perspective of the male victim and sympathize with him more than women.

Study 4

In addition to demonstrating the robustness of the implicit versus explicit bias attribution effect, Study 3 suggested that the extent to which participants take the victim’s or perpetrator’s perspective (and feel sympathy for each of them) may influence judgments of accountability and punishment. Although these processes do not appear to weaken (or enhance) the bias attribution effect on judgments of perpetrator accountability and support for punishment, as originally predicted, they nonetheless hint at a potential role for shared gender group membership in shaping these outcomes. But, it is also possible that gender socialization in general or, rather, experiences as members of a group often targeted for discrimination, leads women to adopt the perspective of, and sympathize with, victims of discrimination more than men, irrespective of whether or not they share a group membership with the victim. Study 4 seeks to examine these possibilities by assessing men and women’s judgments of accountability and support for punishment regarding a female perpetrator of gender discrimination toward a male victim.

Design and Materials

Study 4 had a 2 (Bias Attribution: Explicit vs. Implicit) \times 2 (Participant Gender: Men vs. Women) between-subjects design.

Participants. Through Prolific, 706 people in the United States completed the study in exchange for US\$1.25. Sixteen participants were excluded from analyses: six for not confirming that they answered carefully or accurately and 10 who reported a gender other than “man” or “woman.” Thus, we had a final sample of 690 participants (59.1% female, 69.0% White, $M_{age} = 36.84$, 344 in the explicit condition).

This final sample provided more than 95% power to detect an interaction even assuming a small effect size ($\eta^2_{\text{partial}} = .02$).

Discrimination scenario and bias attribution manipulation. Similar to Studies 2 and 3, participants read about an instance of gender discrimination in the workplace. The discrimination, however, was perpetrated by a female manager (Kathy) against a male employee (Joel). Again, see https://osf.io/tybvq/?view_only=137c871f85f3468eafc263d1324708b0 for the full study materials.

Primary outcome variables. Perceptions of accountability, support for punishment, perspective-taking, and sympathy were all measured as described in Study 3, but adapted for the current context.

Manipulation check. The manipulation check was measured the same as in Studies 2 and 3. In the present study, 111 people failed the manipulation check: 82 people (23.8%) in the explicit condition and 29 people (8.4%) in the implicit condition. In this study, more people in the explicit condition failed the manipulation check, $\chi^2(1, N = 690) = 30.53$, $p < .001$. The results presented here include participants who failed the manipulation check. That said, there are two minor differences on the results when these participants are excluded: (a) the marginal effect of gender on support for punishment becomes nonsignificant, and (b) the effect of bias attribution condition on perpetrator perspective-taking becomes significant. All other main effects and interactions are unchanged as are the mediation analyses.

Other measures. We also included exploratory measures of perceptions of perpetrator intent, the controllability of the perpetrator’s behavior, and the perceived harm to the victim. Analyses of these variables suggested no effects of participant gender and, thus, we report the effects of bias attribution on these variables in the Supplemental Material available online.

Procedure

After providing informed consent, participants read the discrimination scenario followed by the bias attribution manipulation information. They then completed the measures of accountability, punishment, intent, harm, controllability, perspective-taking, and sympathy, in that order. Next, they completed the bias attribution condition manipulation check followed by demographic information (e.g., age, gender, and political orientation). Finally, they reported whether they answered carefully and accurately, were thanked, debriefed, and paid.

Results

Each dependent variable was submitted to a 2 (Bias Attribution: Explicit vs. Implicit) \times 2 (Participant Gender:

Male vs. Female) ANOVA. Mean values and standard deviations for accountability and punishment are provided in Table 2.

Accountability. Replicating the previous studies, analyses revealed that participants in the implicit bias condition held the perpetrator less accountable than those in the explicit bias condition, $F(1, 686) = 45.62, p < .001, \eta^2_{\text{partial}} = .062$. Contrary to the past studies, however, participant gender did not predict perceptions of accountability, $F(1, 686) = 0.49, p = .48, \eta^2_{\text{partial}} = .001$. Importantly, and inconsistent with both the gender socialization and discrimination sensitization hypotheses, women did not hold the female perpetrator in the scenario more accountable for her discrimination than men. Once again, the participant gender by bias attribution interaction was not significant, $F(1, 686) = 0.00, p = .96, \eta^2_{\text{partial}} < .001$.

Punishment. Participants in the implicit bias condition were less supportive of punishing the manager than participants in the explicit bias condition, $F(1, 686) = 50.54, p < .001, \eta^2_{\text{partial}} = .069$. Although it did not reach conventional levels of significance, men were somewhat more supportive of punishing the female perpetrator than were women, $F(1, 686) = 3.20, p = .074, \eta^2_{\text{partial}} = .005$ —an outcome consistent with the predicted role of shared group membership in this work. Similar to all previous studies, participant gender did not moderate the bias attribution effect, $F(1, 686) = 0.01, p = .94, \eta^2_{\text{partial}} < .001$.

Perspective-taking. Also consistent with the putative role of shared gender group membership in this work, rather than gender socialization or sensitization to discrimination, analyses revealed that men ($M = 5.37, SD = 1.40$), who shared a gender group membership with the victim of the discrimination, reported taking the victim's perspective more than women ($M = 4.98, SD = 1.66$), $F(1, 686) = 10.81, p = .001, \eta^2_{\text{partial}} = .016$. And, women ($M = 3.71, SD = 1.75$), who shared a gender group membership with the perpetrator, reported taking her perspective more than men ($M = 3.35, SD = 1.74$), $F(1, 686) = 5.59, p = .018, \eta^2_{\text{partial}} = .008$. Although the effect of bias attribution condition on perspective-taking with the perpetrator was close to conventional levels of statistical significance, $F(1, 686) = 3.10, p = .079, \eta^2_{\text{partial}} = .004$, no other main effects or interactions were significant ($ps > .29$).²

Sympathy. Contrary to predictions, participant gender did not predict feeling sympathy for the male victim, $F(1, 686) = 1.88, p = .17, \eta^2_{\text{partial}} = .003$, or for the female perpetrator, $F(1, 686) = 0.16, p = .69, \eta^2_{\text{partial}} < .001$. There was, however, an unexpected main effect of bias attribution condition on both outcomes. Participants in the implicit bias condition ($M = 4.08, SD = 1.37$) reported feeling *less* sympathy for the victim than those in the explicit bias condition ($M = 4.51,$

$SD = 1.45$), $F(1, 686) = 12.08, p = .001, \eta^2_{\text{partial}} = .017$. Furthermore, those in the implicit bias condition ($M = 2.67, SD = 1.05$) felt *more* sympathy for the perpetrator than those in the explicit bias condition ($M = 2.41, SD = 1.09$), $F(1, 686) = 9.53, p = .002, \eta^2_{\text{partial}} = .014$. The participant gender by bias condition interaction was not significant for either outcome ($ps > .23$).

Mediation analyses. Although participant gender did not have a significant direct effect on either accountability or support for punishment, this is not a necessary condition to test for intervening effects between predictors and outcome variables (Hayes, 2009). Here, we tested the same mediation model from Study 3 to see whether perspective-taking with the perpetrator and victim and/or sympathy for the perpetrator and victim played an intervening role between participant gender and the primary outcome variables: perpetrator accountability and support for punishment.³ Consequently, similar to Study 3, we examined victim and perpetrator perspective-taking (centered), as well as victim and perpetrator sympathy (centered), as mediating processes between participant gender (men = -0.5 , women = 0.5) and the outcome variables, using PROCESS (5,000 bootstrapped resamples, Model 4; Hayes, 2013) with bias attribution condition included as a covariate. As depicted in Figure 2, participant gender significantly predicted victim perspective-taking, $B = -0.39, SE = 0.12, p = .001$, and perpetrator perspective-taking, $B = 0.32, SE = 0.14, p = .018$. Men reported taking the male victim's perspective more than women; whereas women reported taking the female perpetrator's perspective more than men. Furthermore, perpetrator perspective-taking significantly predicted judgments of perpetrator accountability, $B = -0.08, SE = 0.02, p < .001$ (total indirect effect: $B = -0.09, SE = 0.05, 95\% CI = [-.19, .00]$). That is, the more participants took the female perpetrator's perspective, the less they held her accountable. As with the ANOVAs, the direct effect of participant gender on accountability was not significant, $B = 0.04, SE = 0.06, p = .55$ (total effect: $B = -0.05, SE = 0.08, p = .48$).

Support for punishing the perpetrator revealed a similar pattern (total indirect effect: $B = -0.11, SE = 0.06, 95\% CI = [-.22, .00]$). As in the accountability model, women reported taking the female perpetrator's perspective more than men did, $B = 0.32, SE = 0.14, p = .018$, and the tendency to take the perpetrator's perspective was associated with punishing the perpetrator less, $B = -0.08, SE = 0.02, p = .001$. The direct effect of participant gender on punishment was not significant, $B = -0.05, SE = 0.07, p = .47$ (total effect: $B = -0.16, SE = 0.09, p = .074$).

Discussion

Study 4 provided additional evidence that attributing gender discrimination to implicit rather than explicit bias reduces perceptions of perpetrator culpability. Participants who were

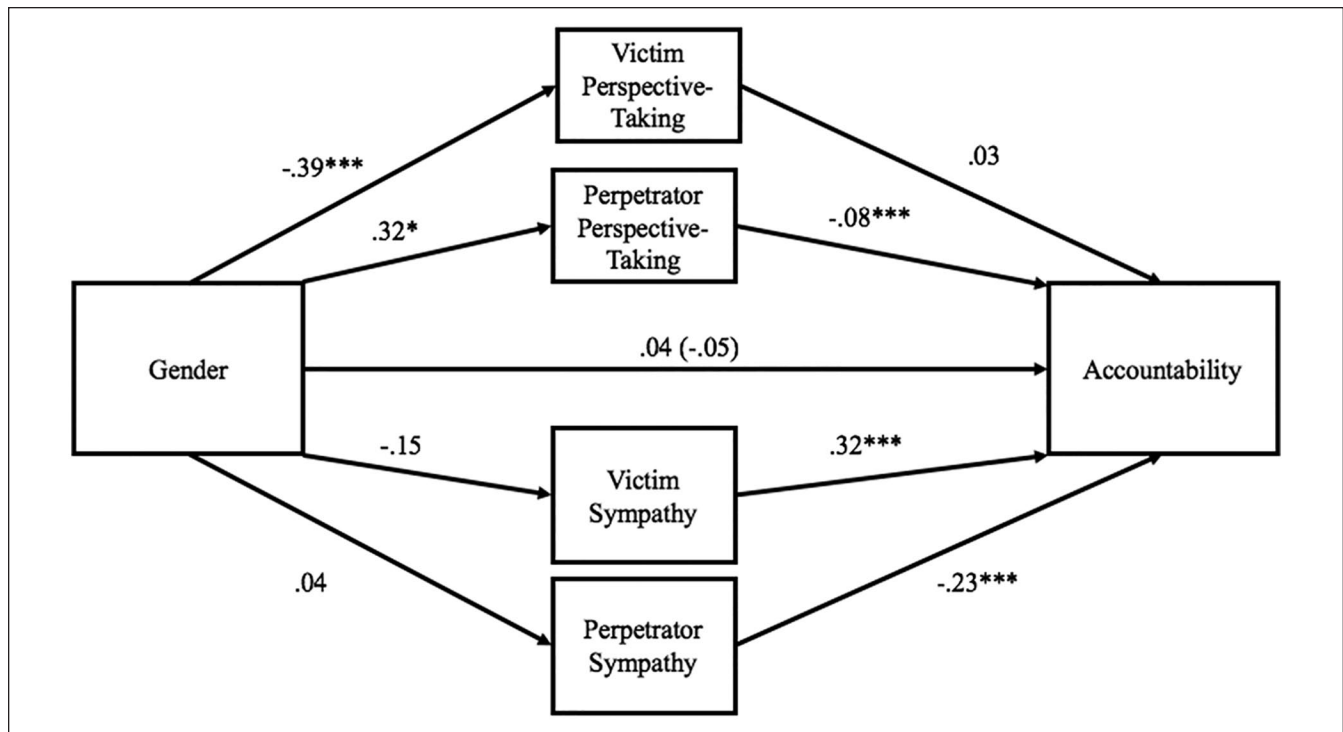


Figure 2. Mediation model depicting the indirect effect of gender on accountability in Study 4.

Note. Bias attribution condition was included as a covariate.

* $p < .05$. ** $p < .01$. *** $p \leq .001$.

told that a female manager engaged in gender discrimination toward a male employee because of her implicit bias held her less accountable and were less supportive of punishing her than participants who were told she discriminated because of explicit bias. Once again, both men and women revealed this bias attribution effect. By reversing the contingencies between participant gender and that of the perpetrator and victim, Study 4 allowed us an opportunity to assess whether shared gender group membership may be playing a role in participants' judgments of perpetrator accountability and support for punishment.

Although the present results were not a full reversal of the past three studies, they are largely inconsistent with the more general gender socialization and/or discrimination sensitization accounts, given that women did not hold the perpetrator of the discrimination in this study more accountable, nor did they support punishing her more, compared with men. Instead, consistent with the putative role of shared group membership, men were somewhat more supportive of punishing the female perpetrator of gender discrimination against a male victim than were women. Furthermore, it was the differential tendency for participants to adopt the perspective of the perpetrator in the scenario that predicted the extent to which participants held the female perpetrator accountable and supported punishing her for the discrimination. In other words, the overall pattern of results suggests that shared gender group membership is playing a role in

these outcomes. To be clear, this does not foreclose on the possibility that women's gender socialization and/or their sensitization to discrimination are also operative and, perhaps, account for the somewhat inconsistent findings across the dependent measures. Indeed, it is possible that the null effect of participant gender on accountability in Study 4 is because women experience a tension between the effects of sharing a relevant group membership with the female perpetrator and sharing the *experience* of discrimination with the male victim (Cortland et al., 2017). Future research is needed to explore these possibilities further.

General Discussion

The present research aimed to examine whether (a) people hold perpetrators less accountable for gender discrimination attributed to implicit compared with explicit bias, and (b) whether this effect would be moderated by participant gender, presumably due to known effects of sharing a relevant group membership with either the perpetrator or victim of the discrimination. Replicating previous work (Cameron et al., 2010; Daumeyer et al., 2019; Redford & Ratliff, 2016), we found consistent evidence that both male (Studies 1–3) and female (Study 4) perpetrators are held less accountable and deemed less worthy of punishment for gender discrimination when it is attributed to their implicit, rather than explicit, attitudes and beliefs. Across these four studies,

however, we found no evidence that sharing a gender (i.e., relevant) group membership with the victim rather than perpetrator of gender discrimination affects the extent to which people hold perpetrators differentially accountable for discrimination attributed to implicit relative to explicit bias. Given that shared group membership could reasonably be expected to shift cognitive and motivational processes relevant to assessing responsibility for harmful acts, including gender discrimination (Branscombe et al., 1999; Major & O'Brien, 2005), the failure of participant gender to moderate the magnitude of the bias attribution effect in this work is quite noteworthy.

Although we found no evidence of moderation of the bias attribution effect by participant gender, women consistently held male perpetrators of gender discrimination against other women more accountable and were more supportive of punishing them than men (Studies 1–3). The mediation analyses in Study 3 revealed that this effect could be explained statistically by women's and men's differential tendency to adopt the perspective of female victim and feel sympathy for the female victim and male perpetrator. When the victim of the discrimination was a man and the perpetrator was a woman (Study 4), a more atypical instance of discrimination, women no longer held the perpetrator more accountable. Once again, however, shared group membership predicted patterns of differential perspective-taking with the perpetrator, which accounted for differences in the extent to which men and women held the female perpetrator accountable and supported punishing her. Taken together, these findings suggest that, whereas sharing a relevant group membership with the perpetrator or target of discrimination appears to influence judgments of accountability and punishment overall, it does not necessarily influence the effect of attributing discrimination to implicit rather than explicit bias on these outcomes.

Implications

The present work builds on a growing body of literature demonstrating that perpetrators are deemed less culpable when their discriminatory behavior is attributed to their implicit rather than explicit biases and beliefs (Cameron et al., 2010; Daumeyer et al., 2019; Redford & Ratliff, 2016). Past work has found that when discriminatory behavior is linked to implicit bias, people infer that it was unintentional and/or uncontrollable, which leads them to hold perpetrators less accountable (Cameron et al., 2010; Onyeador, 2017), presumably because people prioritize intentionality when evaluating wrongdoing (Cushman, 2008; Ginther et al., 2016; Swim et al., 2003). The results of the present work suggest that sharing a relevant group membership with the victim, rather than perpetrator, of the discrimination does not necessarily disrupt this reasoning process. That is, women (who shared a gender group membership with the victim of the discrimination in Studies 1–3) and men (who shared a gender group membership with the victim in Study 4) held

the perpetrators of discrimination against a fellow ingroup member less accountable when the discrimination was attributed to their implicit, rather than explicit, bias. And, importantly, the magnitude with which these participants demonstrated the bias attribution effect was not meaningfully different from that observed among participants who shared a relevant group membership with the perpetrators of the discrimination (i.e., men in Studies 1–3, women in Study 4). Shared group membership, in other words, does not necessarily attenuate the extent to which implicit bias attributions reduce perceived culpability for discrimination.

Although the predicted effect of shared group membership on the magnitude of the bias attribution effect was not observed, sharing a gender with either the perpetrator or victim of the discrimination did appear to influence perceptions of culpability overall. In both Studies 3 and 4, gender predicted differential patterns of perspective-taking with the victim and/or perpetrator, which in turn predicted and partially accounted for differential willingness to hold the perpetrator accountable and punish him or her. Specifically, in Study 3, sympathy toward and perspective-taking with the victim mediated the relationship between participant gender and accountability. Whereas, in Study 4, it was perspective-taking with the perpetrator that mediated the relationship between participant gender and accountability. These findings suggest that shifting who people perspective-take and sympathize with in instances of discrimination, even those attributed to implicit bias, may lead them to rethink the extent to which they should hold the perpetrators accountable (Simon, Moss, & O'Brien, 2019). Future research should investigate this question and, further, whether inducing perspective-taking with the victim might also reduce the magnitude of the bias attribution effect, especially relative to inducing perspective-taking with the perpetrator.

Limitations and Future Directions

Although this work adds to the growing literature on how people reason about discrimination that is attributed to implicit rather than explicit bias, it is not without limitations. Because we chose to focus on gender discrimination in the present work, this limits our ability to generalize to instances of shared group membership based on other dimensions of social identity and in other domains of discrimination (e.g., race and racism). That is, the same patterns may not emerge when Black Americans are making judgments about the accountability of a White American perpetrator of racial discrimination against a Black victim. Indeed, past work suggests that Black and White Americans assess instances of racial discrimination differently (Carter & Murphy, 2015; Corning & Bucchianeri, 2010; Inman & Baron, 1996; Sherman et al., 1983), with race shaping the extent to which people perceive intentionality and harm in instances of racial discrimination (Simon, Moss, & O'Brien, 2019). Thus, sharing a racial group membership with the victims, rather than

perpetrators, of racial discrimination may indeed moderate the extent to which people hold perpetrators less accountable for discrimination that has been attributed to implicit, rather than explicit, bias. Future work is needed to explore this question and we advise against making broad claims about the role of shared group membership in these processes based on the current work.

Examining other domains of discrimination will also afford a better test of the potential role of perspective-taking with victims of discrimination across groups with an intersectional approach. For instance, in Studies 3 and 4, we found evidence that people who shared a gender identity with a victim of gender discrimination tended to take the perspective of the victim more than people who shared a gender identity with the perpetrator. Furthermore, Study 3 provided evidence that this victim-focused approach helped account for people holding the perpetrator more accountable. Would we find similar patterns if the victim of the discrimination was a Black woman or a gay man? People may be more likely to adopt the perspective of prototypical victims of discrimination (e.g., Purdie-Vaughns & Eibach, 2008) and/or only with people with whom they share other relevant identities (e.g., race, sexual orientation). Or, rather, people may be encouraged to think about commonalities between different types of discrimination (e.g., racism, sexism), which, in turn, may increase the extent to which they adopt the perspective of the victim irrespective of whether or not they share a relevant group membership (Cortland et al., 2017). Future work should explore these possibilities both within the domain of gender discrimination and in other domains.

Finally, the current design only explored situations in which participants shared a group membership with either the perpetrator or victim. Future work could explore how people respond to instances of discrimination wherein they share a group membership with both the victim and perpetrator. For example, how might women respond to implicit versus explicit gender discrimination enacted by a female perpetrator toward a female victim? This situation would allow researchers to further probe the extent to which shared group membership (and subsequent perspective-taking/sympathy) versus gender socialization or even general sensitivity toward victims of discrimination shapes perceptions of accountability. In such a case, because women may take the perspective of/sympathize with both the victim and the perpetrator, these variables may not explain their perceptions of accountability and support for punishment.

Conclusion

As implicit bias becomes an increasingly popular explanation for gender discrimination among the general public, it is important for researchers to understand how people reason about the discrimination that is attributed to it. The present work is the first (of which we are aware) to examine this

question and, further, to offer full consideration of whether shared gender group membership shapes the extent to which people hold the perpetrators less accountable for gender discrimination attributed to implicit bias. Indeed, our findings suggest that sharing a gender group membership with the victims of gender discrimination does not reduce, much less eliminate, the bias attribution effect on perpetrator accountability or support for punishment. Given the robustness of the bias attribution effect to reduce perpetrator accountability and punishment, then, the present work suggests a need to be far more thoughtful about when and why we attribute specific instances of discrimination to implicit rather than explicit bias. Misattributing the cause of discrimination, in other words, may prove especially detrimental to efforts to combat, if not root out, sexism in the workplace and society at large.

Authors' Note

Ivuoma N. Onyeador is now affiliated with Northwestern University, Evanston, IL, USA.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by an NSF Social, Behavioral, and Economic Sciences (SBE) Postdoctoral Research Fellowship (#1809370) awarded to the second author as well as by an NSF Behavioral and Cognitive Sciences (BCS) Grant (#1941651) awarded to the third author.

ORCID iDs

Natalie M. Daumeyer  <https://orcid.org/0000-0002-9223-0684>

Ivuoma N. Onyeador  <https://orcid.org/0000-0002-4272-8572>

Supplemental Material

Supplemental material is available online with this article.

Notes

1. Overall, 31 participants (10.6%) listed an occupation in a science, technology, engineering, and mathematics (STEM) field or company. When included as a covariate, working in STEM did not significantly affect the outcome variables ($ps > .14$). Thus, it will not be discussed further.
2. When participants who failed the manipulation check are excluded, the bias attribution effect on perpetrator perspective-taking becomes significant with participants in the implicit bias condition ($M = 3.68$, $SD = 1.70$) taking the perspective of the perpetrator more than those in the explicit condition ($M = 3.30$, $SD = 1.72$), $F(1, 575) = 4.07$, $p = .044$, $\eta^2_{\text{partial}} = .007$.

3. Although perspective-taking and sympathy were measured after accountability and punishment in this study, we used the same mediation model as Study 3, given our interest in exploring whether the pattern of results that emerged there would also emerge in the present study with a female perpetrator and male victim. Because there was no direct effect of participant gender on accountability in Study 4, we are confident that perceived accountability does not mediate the relationship between participant gender and perspective-taking/sympathy.

References

- Altmire, J. (2019, May 10). Gender bias is real—This is how 2020 candidates can overcome it. *The Hill*. <https://thehill.com/opinion/campaign/442414-gender-bias-is-real-this-is-how-2020-candidates-can-overcome-it>
- Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 35–58). Blackwell.
- Brewer, M. B., & Kramer, R. M. (1985). The psychology of intergroup attitudes and behavior. *Annual Review of Psychology*, *36*, 219–243.
- Cameron, C. D., Payne, B. K., & Knobe, J. (2010). Do theories of implicit race bias change moral judgments? *Social Justice Research*, *23*(4), 272–289. <https://doi.org/10.1007/s11211-010-0118-z>
- Carter, E. R., & Murphy, M. C. (2015). Group-based differences in perceptions of racism: What counts, to whom, and why? *Social and Personality Psychology Compass*, *9*(6), 269–280. <https://doi.org/10.1111/spc3.12181>
- Corning, A. F., & Bucchianeri, M. M. (2010). Perceiving racism in ambiguous situations: Who relies on easy-to-use information? *The Journal of Social Psychology*, *150*(3), 258–277. <https://doi.org/10.1080/00224540903365398>
- Cortland, C. I., Craig, M. A., Shapiro, J. R., Richeson, J. A., Neel, R., & Goldstein, N. J. (2017). Solidarity through shared disadvantage: Highlighting shared experiences of discrimination improves relations between stigmatized groups. *Journal of Personality and Social Psychology*, *113*(4), 547–567. <https://doi.org/10.1037/pspi0000100>
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–380. <https://doi.org/10.1016/j.cognition.2008.03.006>
- Daumeyer, N. M., Onyeador, I. N., Brown, X., & Richeson, J. A. (2019). Consequences of attributing discrimination to implicit vs. Explicit bias. *Journal of Experimental Social Psychology*, *84*, Article 103812. <https://doi.org/10.1016/j.jesp.2019.04.010>
- Ginther, M. R., Bonnie, R. J., Hoffman, M. B., Shen, F. X., Simons, K. W., Jones, O. D., & Marois, R. (2016). Parsing the behavioral and brain mechanisms of third-party punishment. *The Journal of Neuroscience*, *36*(36), 9420–9434. <https://doi.org/10.1523/JNEUROSCI.4499-15.2016>
- Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491–512. <https://doi.org/10.1037/0022-3514.70.3.491>
- Guterk, B. A., Cohen, A. G., & Tsui, A. (1996). Reactions to perceived sex discrimination. *Human Relations*, *49*(6), 791–813.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408–420. <https://doi.org/10.1080/03637750903310360>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.
- Hess, U., Senécal, S., Kirouac, G., Herrera, P., Philippot, P., & Kleck, R. E. (2000). Emotional expressivity in men and women: Stereotypes and self-perceptions. *Cognition & Emotion*, *14*(5), 609–642. <https://doi.org/10.1080/02699930050117648>
- Inman, M. L., & Baron, R. S. (1996). Influence of prototypes on perceptions of prejudice. *Journal of Personality and Social Psychology*, *70*, 727–739.
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). Academic Press.
- Lithari, C., Frantzidis, C. A., Papadelis, C., Vivas, A. B., Klados, M. A., Kourtidou-Papadeli, C., Pappas, C., Ioannides, A. A., & Bamidis, P. D. (2010). Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. *Brain Topography*, *23*(1), 27–40. <https://doi.org/10.1007/s10548-009-0130-5>
- Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442. <https://doi.org/10.3758/s13428-016-0727-z>
- Major, B., & O'Brien, L. T. (2005). The social psychology of stigma. *Annual Review of Psychology*, *56*(1), 393–421. <https://doi.org/10.1146/annurev.psych.56.091103.070137>
- Major, B., Quinton, W. J., & McCoy, S. K. (2002). Antecedents and consequences of attributions to discrimination: Theoretical and empirical advances. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 251–330). Elsevier. [https://doi.org/10.1016/S0065-2601\(02\)80007-7](https://doi.org/10.1016/S0065-2601(02)80007-7)
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *The Journal of Psychology*, *25*(1), 147–186. <https://doi.org/10.1080/1047840X.2014.877340>
- Manne, K. (2019, April 11). It's the sexism, stupid. *Politico*. <https://www.politico.com/magazine/story/2019/04/11/its-the-sexism-stupid-226620>
- McNamara, R. A., Willard, A. K., Norenzayan, A., & Henrich, J. (2019). Weighing outcome vs. intent across societies: How cultural models of mind shape moral reasoning. *Cognition*, *182*, 95–108. <https://doi.org/10.1016/j.cognition.2018.09.008>
- Monroe, A. E., & Malle, B. F. (2018). People systematically update moral judgments of blame. *Journal of Personality and Social Psychology*, *116*, 215–236. <http://dx.doi.org/10.1037/pspa0000137>
- Mundy, L. (2017, April). Why is Silicon Valley so awful to women? *The Atlantic*. https://www.theatlantic.com/magazine/archive/2017/04/why-is-silicon-valley-so-awful-to-women/517788/?utm_source=atlwt
- Nadler, J., & McDonnell, M.-H. (2011). Moral character, motive, and the psychology of blame. *Cornell Law Review*, *97*, 255–304.
- Onyeador, I. N. (2017). *Presumed unintentional: The ironic effects of implicit bias framing on Whites' perceptions of discrimination* [Unpublished doctoral dissertation]. University of California.
- Parker, K., & Funk, C. (2017, December 14). *Gender discrimination comes in many forms for today's working women*. Pew Research

- Center. <https://www.pewresearch.org/fact-tank/2017/12/14/gender-discrimination-comes-in-many-forms-for-todays-working-women/>
- PayScale. (2019). *The state of the gender pay gap 2019*. <https://www.payscale.com/data/gender-pay-gap>
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811–832.
- Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles*, *59*(5–6), 377–391. <https://doi.org/10.1007/s11199-008-9424-4>
- Redford, L., & Ratliff, K. A. (2016). Perceived moral responsibility for attitude-based discrimination. *British Journal of Social Psychology*, *55*(2), 279–296. <https://doi.org/10.1111/bjso.12123>
- Sherman, M. F., Smith, R. J., & Sherman, N. C. (1983). Racial and gender differences in perceptions of fairness: When race is involved in a job promotion. *Perceptual and Motor Skills*, *57*(3), 719–728. <https://doi.org/10.2466/pms.1983.57.3.719>
- Simon, S., Moss, A. J., & O'Brien, L. T. (2019). Pick your perspective: Racial group membership and judgments of intent, harm, and discrimination. *Group Processes & Intergroup Relations*, *22*, 215–232. <https://doi.org/10.1177/1368430217735576>
- Stürmer, S., Snyder, M., & Omoto, A. M. (2005). Prosocial emotions and helping: The moderating role of group membership. *Journal of Personality and Social Psychology*, *88*(3), 532–546. <https://doi.org/10.1037/0022-3514.88.3.532>
- Swim, J. K., Scott, E. D., Sechrist, G. B., Campbell, B., & Stangor, C. (2003). The role of intent and harm in judgments of prejudice and discrimination. *Journal of Personality and Social Psychology*, *84*(5), 944–959. <https://doi.org/10.1037/0022-3514.84.5.944>
- Tanis, M., & Postmes, T. (2005). A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour. *European Journal of Social Psychology*, *35*(3), 413–424. <https://doi.org/10.1002/ejsp.256>
- Uhlmann, E. L., & Nosek, B. A. (2012). My culture made me do it: Lay theories of responsibility for automatic prejudice. *Social Psychology*, *43*(2), 108–113. <https://doi.org/10.1027/1864-9335/a000089>
- Weisbuch, M., & Ambady, N. (2008). Affective divergence: Automatic responses to others' emotions depend on group membership. *Journal of Personality and Social Psychology*, *95*(5), 1063–1079. <https://doi.org/10.1037/a0011993>
- Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership modulates empathic neural responses. *Journal of Neuroscience*, *29*(26), 8525–8529. <https://doi.org/10.1523/JNEUROSCI.2418-09.2009>
- Ziv, S. (2016, January 15). The “Elephant in the Valley”: A new survey looks at the experiences of women working in tech. *Newsweek*. <https://www.newsweek.com/silicon-valley-women-tech-survey-416206>