

## Man up and take it: Gender bias in moral typecasting

Tania Reynolds<sup>a,b,\*</sup>, Chuck Howard<sup>c</sup>, Hallgeir Sjøstad<sup>d</sup>, Luke Zhu<sup>e</sup>, Tyler G. Okimoto<sup>f</sup>, Roy F. Baumeister<sup>f</sup>, Karl Aquino<sup>c</sup>, JongHan Kim<sup>g</sup>

<sup>a</sup> University of New Mexico, Psychology Department, United States

<sup>b</sup> Indiana University Bloomington, Kinsey Institute and Gender Studies Department, United States

<sup>c</sup> University of British Columbia, Sauder School of Business, Canada

<sup>d</sup> Norwegian School of Economics and Center for Applied Research at NHH, Norway

<sup>e</sup> York University, Schulich School of Business, Canada

<sup>f</sup> The University of Queensland, Australia

<sup>g</sup> Coastal Carolina University, Department of Psychology, United States

### ARTICLE INFO

#### Keywords:

Moral typecasting  
Gender stereotypes  
Victimization  
Punishment  
Justice  
Morality  
Bias

### ABSTRACT

Informed by moral typecasting theory, we predicted a gender bias in harm evaluation, such that women are more easily categorized as victims and men as perpetrators. Study 1 participants assumed a harmed target was female (versus male), but especially when labeled ‘victim’. Study 2 participants perceived animated shapes perpetuating harm as male and victimized shapes as female. Study 3 participants assumed a female employee claiming harassment was more of a victim than a male employee making identical claims. Female victims were expected to experience more pain from an ambiguous joke and male perpetrators were prescribed harsher punishments (Study 4). Managers were perceived as less moral when firing female (versus male) employees (Study 5). The possibility of gender discrimination intensified the cognitive link between women and victimhood (Study 6). Across six studies in four countries ( $N = 3,137$ ), harm evaluations were systematically swayed by targets’ gender, suggesting a gender bias in moral typecasting.

### 1. Introduction

The general public, shareholders, media pundits, and organizational scholars are seemingly united in demanding managers treat employees impartially. When managers are suspected of acting otherwise, as when they hire or promote contingently based on irrelevant factors such as employees’ race or gender, observers are quick to charge them with discrimination. Researchers have devoted much effort to documenting how gender stereotypes impair evaluators’ ability to view women as competent leaders, hindering ascension to the highest levels of organizations (Eagly & Karau, 2002; Heilman, 2001). Beyond demonstrating impartiality when making selection and promotion decisions, managers are also increasingly expected to demonstrate impartial concern about the physical, psychological, and emotional well-being of employees. Yet, despite exhortations to act without prejudice, in practice, it is difficult to achieve these lofty goals. Even a cursory reading of the organizational justice literature reveals managers routinely violate impartiality, which many philosophers and management scholars consider a prerequisite for logical reasoning, the proper administration of justice,

and moral virtue. For this reason, it is a pressing concern for organizational scholars to document how managerial decision-making can be swayed from impartiality.

Decades of research reveal the failure to adopt an impartial spectator’s perspective is a seemingly intractable feature of human psychology. Impartiality is compromised by unconscious forces that produce systematic biases in judgment. We define bias as a systematic deviation from rational consistency, whereby judgment is influenced by factors irrelevant to the ostensible goal (Thaler, 2015). Researchers have identified a host of automatic cognitive processes constructed by natural selection to facilitate rapid, reflexive decision making (Kahneman, 2011; Haidt, 2012; Hauser, 2006). When evaluating moral events, which we define as situations involving harm, people’s reliance on mental shortcuts can lead them to exhibit what Bazerman and Tenbrunsel (2011) refer to as ethical “blind spots”, even among those of professed goodwill.

In this paper, we examine one type of mental shortcut, *moral typecasting* (Gray & Wegner, 2009), to investigate whether observers exhibit a gender bias in their assessments of moral events. The moral

\* Corresponding author at: Kinsey Institute and Department of Gender Studies, Indiana University Bloomington, 430 Lindley Hall, 150 S. Woodlawn Ave., Bloomington, IN 47405, United States.

E-mail address: [tareyn@iu.edu](mailto:tareyn@iu.edu) (T. Reynolds).

<https://doi.org/10.1016/j.obhdp.2020.05.002>

Received 17 January 2018; Received in revised form 15 May 2020; Accepted 18 May 2020

0749-5978/ Published by Elsevier Inc.

typecasting framework proposes that humans instinctively perceive moral behavior through a cognitive template, in which they cast parties into the dyadic roles of intentional “agent” or suffering “patient.” We hypothesize that gender stereotypes and base rates of harm facilitate categorizing women into the role of suffering patient and men into the role of perpetrating agent, which leads decision-makers to exhibit systematic biases in applying this template. When evaluating harm, perceivers will be swayed by the involved targets’ gender to more readily categorize women as victims and men as perpetrators than the converse. As a result, managers might unknowingly exhibit ethical blind spots by more readily detecting and more emphatically responding to female victimization and male perpetration than male victimization or female perpetration.

Previous studies demonstrate that moral typecasting is a reflexively employed cognitive prototype, but less research has examined whether it is subject to bias (but see [FeldmanHall et al., 2016](#)). If our hypothesis is supported, it would suggest victims of harm receive differential levels of concern, support, and retributive punishment due to factors beyond their control. In their moral calculus of harm suffered by male or female employees, managers may not weigh these outcomes impartially. We predict that perpetrators of harm will be more easily exculpated for their misdeeds if they do not conform heuristically to the perpetrator role or if their targets do not conform heuristically to the victim role. This gender bias in moral typecasting could thus impair managers’ ability to act justly when adjudicating workplace disputes, responding to accusations of harm, and deciding on appropriate remedial action. This bias may also affect policy-makers and legislators who must decide which harms warrant attention, resources, and intervention. For these reasons, our research has implications for the well-being of millions around the world.

## 2. Theoretical background

### 2.1. Moral typecasting

The moral typecasting hypothesis contends that observers perceive and interpret moral actions and events through a dyadic template ([Gray & Wegner, 2009](#)), in which they apply a cognitive schema that casts involved persons into the roles of intentional “agent” or suffering “patient.” This schema is a prototype observers apply when one or more parties appears to experience harm. Observers apply this dyadic template to categorize actors depending on the features of the situation and characteristics of those actors. Moral typecasting theory further posits that the roles of agent and patient are mutually exclusive, meaning that when observers perceive an individual as an agent, they are less likely to view that same individual as a patient, and vice versa.

The application of the dyadic template has practical consequences because the roles of agent and patient evoke divergent emotional responses and moral judgments from observers. Those assigned to the patient category are expected to experience more pain and suffering, thereby eliciting greater sympathy, compared to those typecast into the agentic role ([Gray & Wegner, 2009](#)). Patients can therefore be construed as “victims” of harm whereas agents are construed as “perpetrators” who are responsible, intentional, and deserving of blame and punishment ([Gray, Gray, & Wegner, 2007](#); [Gray & Wegner, 2011](#)). These categorizations influence how observers subsequently evaluate and respond to the actors assigned to these roles.

Although Gray and colleagues argue the dyadic template is a general schema employed for evaluating moral actions, they also suggest that “features that make harm salient should decrease the moral acceptability of actions” (p. 209; [Gray, Waytz, & Young, 2012](#)). Put another way, observers may not always respond to harm evenhandedly and instead exhibit asymmetric patterns of outrage or sympathy when a victim’s harm is particularly salient. If moral responses can be swayed by situational factors such as harm salience, it is possible asymmetries might also result from the degree to which targets fit the prototypical

examples of agent or patient. This relative degree of prototypicality may influence the cognitive ease with which observers categorize individuals into each role and recognize suffering. We contend that gender carries a multitude of stereotypical associations relevant for assigning people into agent and patient roles. We hypothesize that gender stereotypes facilitate women’s being typecast as victims and men’s being typecast as perpetrators. If so, gendered typecasting should result in a greater ease of detecting women’s (versus men’s) victimization and men’s (versus women’s) perpetration. This prototypicality should shape moral judgments, such that women’s suffering evokes greater sympathy and men’s perpetration harsher punishments.

### 2.2. Men as agents

Throughout history and still today, men have been perceived as more agentic than women. Explicit notions of masculinity include traits such as dominance, individualism, force, self-sufficiency, and ambition ([Bem, 1974](#); [Heilman, 2001](#); [Rudman & Kilianski, 2000](#)). People also hold these beliefs at subconscious levels and implicitly associate men more closely with power than women ([Rudman, Greenwald, & McGhee, 2001](#)). Men are more likely to view themselves as agentic than are women, endorsing self-descriptions such as active, independent and decisive ([Abele, 2003](#)). Across cultures, men more strongly value power, achievement, and self-direction than do women, suggesting men may more ardently strive for agentic roles ([Schwartz & Rubel, 2005](#)). Cross-cultural behavioral data demonstrate that men are more physically active and physically aggressive than women ([Archer, 2004](#); [Bauman et al., 2012](#)), providing empirical support for gendered social expectations.

Gender stereotypes mirror these discrepancies in behavior, as masculinity is often characterized by perceptions of activity and aggression ([Bem, 1974](#); [Bosson, Vandello, Burnaford, Weaver, & Arzu Wasti, 2009](#)). Implicitly, individuals associate men with threat, violence, destruction, and anger—associations congruent with the perpetrator role ([Rudman & Goodwin, 2004](#); [Rudman et al., 2001](#)). Men’s morphology may strengthen these associations as their bodies have higher proportions of lean muscle mass on average than do women’s ([Lassek & Gaulin, 2009](#)). People perceive muscular, compared to leaner, individuals as less vulnerable, experience less pity in response to their suffering, and are less motivated to protect them ([Dijker, 2001](#)). These patterns suggest people around the world perceive men as more agentic and aggressive than women—traits congruent with the perpetrator role.

If men are more readily typecast as agents, then this categorization should have consequences for third parties’ moral responses. Because the dyadic template roles are mutually exclusive, men’s association with agency should impair their being typecast as passive victims ([Gray & Wegner, 2009](#)). For example, agency predicts perceptions of control, which increases perceptions of responsibility ([Smith & Ellsworth, 1985](#)) and attributions of blame ([Gray et al., 2007](#); [Gray & Wegner, 2011](#)). Perceivers may blame men more than women for whatever misfortunes befall them and feel less motivated to alleviate the consequent suffering ([Cappelen, Falch, & Tungodden, 2017](#); [Weiner, 1980](#)). When targets are perceived as responsible for harm, they evoke less sympathy and aid, as well as more anger from observers ([Rudolph, Roesch, Greitemeyer, & Weiner, 2004](#)). To the degree then that men are viewed as more agentic than women, men should also be assumed to have greater control and responsibility in harm-related contexts. If so, men’s suffering may evoke less sympathy than equivalent suffering by women. Indeed, people expect men to have higher tolerance for pain than women, are more willing to administer shocks to men, and preferentially sacrifice men in trolley dilemmas ([FeldmanHall et al., 2016](#)). This relative inability to perceive men as victims should not only reduce observers’ sympathetic responses to men’s suffering, but should also impair men’s ability to request others reduce or recompense it because such demands are inappropriate for those typecast as perpetrators.

### 2.3. Women as patients

In contrast to male gender stereotypes, female stereotypes traditionally depict women as possessing characteristics consistent with the patient role. The social construct of femininity is associated with traits such as tender, yielding, gentle, and warm (Bem, 1974; Heilman, 2001; Rudman & Goodwin, 2004; Rudman & Kilianski, 2000). These perceptions are consistent with the “women are wonderful” effect, whereby women are viewed favorably because they are assumed to be warm, but not agentic (Eagly & Mladinic, 1989, 1994). Perceptions of high warmth and low agency produce increased pity from social partners (Cuddy, Fiske, & Glick, 2007). Furthermore, those who perceive women as warm but not agentic tend to hold more favorable explicit and implicit attitudes towards women (Rudman & Goodwin, 2004; Rudman & Kilianski, 2000).

Compared to men, adult women are perceived as more child-like, vulnerable, and pain sensitive, and thus evoke more pity and protective inclinations (Bem, 1974; Dijker, 2001, 2010; FeldmanHall et al., 2016). Some portion of women’s greater perceived vulnerability may stem from their facial morphology. On average, women’s faces possess more cues of neoteny or juvenility, including large eyes, small chins, and small noses (Cunningham, Roberts, Barbee, Druen, & Wu, 1995; Tanner, 1978). These facial features are often viewed as cues of naïveté, and they prompt warm and helpful responses from social partners (Berry & McArthur, 1985; Cunningham, 1986; Keating, Randall, Kendrick, & Gutshall, 2003). Such prosocial responses to neotenus features likely stemmed from evolutionary pressures to protect and care for vulnerable offspring (Berry & McArthur, 1985; Bowlby, 1969).

From an evolutionary perspective, individuals may feel more strongly motivated to protect women than men from harm because women set the upper limit on reproduction (Burnstein, Crandall, & Kitayama, 1994). Women are responsible for pregnancy, childbirth, and lactation, and therefore contribute substantially to reproduction. Consider a social group comprised of only three women, but twenty men. This group’s total reproductive success is limited to the few children the three women could bear. Another group comprised of twenty women and only three men could substantially out-reproduce the former. The discrepancy in men’s and women’s reproductive contribution may have favored motivations to protect women from harm. Indeed, people feel a stronger motivation to help women over men, but this gender bias disappears when considering toddlers or elderly individuals, life stages when females are not fertile (Burnstein et al., 1994; Dijker, 2001, 2010). This particular pattern suggests that the preference to protect women over men may stem from evolutionary pressures to insulate reproductively valuable individuals from harm.

If managers and other observers view women as more vulnerable than men, they should find it cognitively easier to typecast women as patients, resulting in a greater likelihood of detecting women’s victimization, and perhaps, stronger motivations to alleviate female suffering. Perceptions of a target’s vulnerability predict pity, concern and moral anger (towards the harm-doer) in response to that target’s suffering (Dijker, 2001, 2010). Researchers have argued that such compassion reflects evolutionary motivations to protect and aid the weak (Goetz, Keltner, Simon-Thomas, 2010; Haidt, 2003; Krebs, 2008). Not only do perceptions of victimhood and feelings of sympathy increase motivation to help sufferers, they also decrease blame ascriptions and punitive motivations toward sufferers (Gray & Wegner, 2011; Rudolph et al., 2004). This pattern suggests if third parties more readily perceive women (vs. men) as victims, female (vs. male) suffering will more easily be recognized and evoke greater sympathy. Likewise, female (vs. male) perpetration of harm will be less easily recognized and more readily excused.

The preceding arguments lead us to hypothesize that individuals will exhibit a biased application of moral typecasting, such that women are more readily typecast as victims, while men are more readily typecast as perpetrators. It is worth noting that this bias is not

independent from true base rates of harm or gender stereotypes. Rather, we contend that harm base rates, such as the greater male perpetration of physical aggression and workplace bullying (Archer, 2004; Workplace Bullying Institute, 2017), contribute to this application of moral typecasting, such that men more closely fit the cognitive prototype of perpetrator. Similarly, gender stereotypes of women as tender and men as dominant (Bem, 1974) likely exacerbate the cognitive expectation of women as victims and men as perpetrators. Importantly, if the cognitive prototypes of victims and perpetrators are gendered, evaluators may not respond to equivalent harms with similar concern or condemnation, even when men and women are involved in identical situations.

The patterns described above suggest female employees, more so than male, will be the beneficiaries of certain types of moral responses including greater recognition of and sympathy towards their suffering. For example, in ambiguous workplace conflicts where both parties claim victimhood, male employees, on average, will more likely be cast as perpetrators, rather than sufferers of harm. If so, male employees may receive disproportionately harsher punishments and disproportionately lower recompense or forbearance than female employees. The main goal of the current investigation was to examine this possibility across various contexts involving men and women as potential agents/perpetrators and patients/victims of harm.

### 2.4. Participant gender and asymmetric moral typecasting

Beyond testing for a general pattern of biased moral typecasting, we also examined whether male and female observers would be equally likely to demonstrate this bias. Competing hypotheses could be generated. On one hand, men score higher than women in benevolent sexism, as measured by items such as ‘women should be cherished and protected by men’, suggesting men might be especially inclined to cast women into the victim role (Glick & Fiske, 1996; Rudman & Kilianski, 2000). On the other hand, women show stronger in-group biases than men, such that they exhibit greater implicit and explicit favoritism toward women, whereas men exhibit more gender-neutral attitudes (Cappelen et al., 2017; Nosek & Banaji, 2002; Richeson & Ambady, 2001; Rudman & Goodwin, 2004). Furthermore, women report experiencing greater prejudice directed at their gender group than do men across a variety of social contexts (Kobrynowicz & Branscombe, 1997; Schmitt, Branscombe, Kobrynowicz, & Owen, 2002). Such experiences of collective victimization may suggest women detect female victimization across situations, and therefore more readily typecast women (vs. men) into the patient role. We tested which of these competing hypotheses would be supported.

## 3. Overview of studies

Study 1 examined whether the cognitive prototype of victim is female rather than male, by testing whether participants assumed a gender-unspecified target of workplace harm was a woman. Study 1 tested whether this gendered expectation would be amplified when harm was made salient by describing actors as ‘victim’ and ‘perpetrator’. Study 2 employed a cross-cultural sample to examine whether participants would exhibit a biased application of moral typecasting when evaluating animated triangles’ interactions, stripped of humanizing characteristics. Study 3 investigated moral typecasting in a workplace dispute where victimization was ambiguous, and included a search for a boundary condition of the gender bias (a neurologically-impaired individual). Study 4 manipulated the genders of both a perpetrator and target of an ambiguous joke to examine downstream consequences of moral typecasting, such as expectations about suffering and willingness to punish the alleged perpetrator. Study 5 tested whether the biased application of moral typecasting extends to harm suffered by groups by examining whether fired female employees are assumed to suffer more than a group of fired male employees. Participants

also evaluated the morality of the managerial team who reached this decision and made inferences about the team's gender composition. Last, Study 6 sought to rule out an alternative explanation to biased application of moral typecasting: possible gender discrimination. Participants evaluated a scenario whereby managers fired either male or female employees from a male-dominated (chemical manufacturing) or female-dominated (nursing) occupation to examine whether cues suggesting gender discrimination either explain or amplify the gender bias in moral typecasting.

Across all six studies on gender-based moral typecasting ( $N = 3,137$ ), we also explored whether male or female participants would exhibit larger asymmetries in their evaluations of harm. We report all measures, conditions, and exclusions in all studies, and always followed APA ethical guidelines.

#### 4. Study 1

If women more closely match the cognitive prototype of victim than do men, then third parties should be more likely to assume a gender-unspecified target who experiences harm (i.e., a victim) is a woman than a man. This pattern should be most pronounced when harm is made salient by explicitly labeling actors as 'victim' and 'perpetrator' as opposed to using more neutral terms (i.e., party). Furthermore, if women are more easily typecast as victims, their harm should evoke more tender responses than men who experience the same harm.

Study 1 participants read vignettes depicting workplace harm, in which the labeling of the targets was experimentally manipulated (IV-1), such that they were described either as 'victim/perpetrator' or 'Party A/Party B'. Within the vignettes, we also experimentally manipulated the gender of the offender (IV-2) to examine whether perpetrator features shift assumptions of victim gender. Participants indicated whether they believed the harmed individual was a man or woman, along with their assessments of the target. We predicted people would assume a female victim, but especially when the targets were labeled 'perpetrator/victim'. Furthermore, we hypothesized that those who assumed a female (versus male) victim would perceive the harm as less deserved and evaluate her more favorably.

#### 4.1. Method

##### 4.1.1. Participants and design.

Three hundred American residents (48% female,  $M_{\text{age}} = 36.7$  years) participated in an online survey through Amazon's MTurk. Participants completed a comprehension check immediately after instructions to verify engagement; the survey was automatically terminated if respondents failed this check, and data collection continued until 300 valid cases were reached. Participants were randomly assigned to the cells of a 2 (offender gender)  $\times$  2 (target labeling: victim/perpetrator vs. Party A/B) between-groups design.

##### 4.1.2. Procedure

After providing consent and basic demographic information, participants read a scenario depicting workplace conflict written from the perspective of the victim. The instructions experimentally manipulated the labeling of the involved individuals (parentheses indicate alternate condition):

*"We are seeking your views about the victim and the perpetrator (major parties) in formal complaints. You will be presented with a direct quote from an actual victim (employee). As you read the case, please consider your thoughts and feelings about both the perpetrator and the victim (parties) of the conflict."*

Next, participants read one of three scenarios, randomly varied to improve the generalizability of the results. Scenario 1 depicted a senior surgeon bullying a surgery trainee in the operating room to the point that the trainee developed suicidal tendencies and depression. Scenario

2 depicted a retail manager forcing an employee to perform tasks that aggravated the employee's preexisting medical conditions, causing the employee to ultimately transfer and take unpaid medical leave. Scenario 3 depicted a postal worker who verbally abused a coworker and filed a false harassment claim to get the coworker fired, leading the coworker to develop an anxiety disorder. Importantly, the scenarios experimentally manipulated the gender of the offender, but avoided gender pronouns when describing the victim.

##### 4.1.3. Dependent measures

After reading the scenario, participants completed a series of items assessing perceptions of the involved targets. (See Appendix for additional measures, which were outside the scope of the primary hypotheses).

**4.1.3.1. Perceived victim gender.** Participants responded to a dichotomous item: "To the best of your recollection, was [the victim / Party A] in this case male or female?"

**4.1.3.2. Victim deservingness.** Participants responded to five ad hoc questions assessing victim deservingness ( $\alpha = 0.81$ ) using a 7-point scale (1 = not at all, 7 = very much): 1) *Perhaps [the target] deserved what happened;* 2) *It seems clear that this ought to have happened to [the target];* 3) *[The target] may, in some way, have deserved what happened in this situation;* 4) *[The target]'s behavior in no way warrants what occurred (reverse);* 5) *It is possible that [the target] was partly responsible for his/her suffering.*

**4.1.3.3. Affective reactions.** Participants provided their feelings toward the victim on a 10-item attitude scale ( $\alpha = 0.88$ ; Philpot & Hornsey, 2008). Participants indicated the degree to which they felt: *positive, happy, warm, good, goodwill, negative[rc], cold[rc], angry[rc], bitter[rc], resentment[rc]* on 7-point scales (1 = not at all, 7 = very much).

**4.1.3.4. Moral character judgments.** Participants rated the victim's moral character on 7-point scales (1 = not at all, 7 = very much): *moral, decent, of good quality, honorable, worthy of respect, immoral[rc], appalling[rc], malicious[rc], and worthless[rc]* ( $\alpha = 0.77$ ; Philpot & Hornsey, 2008).

#### 4.2. Results

##### 4.2.1. Female victim assumption

First, we examined participants' victim gender assumptions by testing the ratio of observed versus expected (50:50) gender categorization. A chi-square analysis indicated participants generally assumed a female victim [76% (72:228)  $\chi^2(1) = 81.12, p < .001$ ], supporting the predicted cognitive link between women and victimhood.

##### 4.2.2. Female victim assumption is amplified by victim/offender labels

Next, we examined whether expectations of the harmed target's gender differed based on targets' labeling through a logistic regression analysis that simultaneously considered labeling, offender gender, their interaction, and control variables (scenario, participant gender). See Table 1 for full results. Step 1 examined the effect of labeling on victim gender assumptions. Supporting hypotheses, participants were more likely to assume a female victim when the targets were labeled perpetrator/victim compared to party A/B [ $B = 0.373, \text{Wald}(1) = 7.154, p = .007$ ; odds ratio = 1.452]. Step 2 entered offender gender and the offender gender  $\times$  labeling interaction. Labeling remained a significant predictor of assumed victim gender [ $B = 0.369, \text{Wald}(1) = 6.037, p = .014$ ; odds ratio = 1.446]. Offender gender also significantly predicted victim gender, such that participants were more likely to assume a female victim when the offender was female [ $B = 0.573, \text{Wald}(1) = 6.037, p < .001$ ; odds ratio = 1.774]. These two main effects were not qualified by a significant offender gender  $\times$  labeling

**Table 1**  
Results of study 1's logistic regression.

	Predictor	B	SE	Wald	Odds ratio
Step 1	Labeling	0.373	0.140	7.154	1.452*
Step 2	Labeling	0.369	0.150	6.037	1.446*
	Offender Gender	0.573	0.150	14.605	1.774*
	Labeling × OffGender	-0.087	0.150	0.337	0.917
Step 3	Labeling	0.384	0.155	6.125	1.469*
	Offender Gender	0.609	0.156	15.262	1.838*
	Labeling × OffGender	-0.146	0.156	0.874	0.864
	Scene 2 dummy	-1.048	0.408	6.600	0.351*
	Scene 3 dummy	-1.628	0.395	16.970	0.196*
	Participant Gender	0.520	0.299	3.029	1.682†

\*  $p < .05$ .

†  $p < .10$ .

interaction [B = -0.087, Wald(1) = 0.34,  $p = .562$ ; odds ratio = 0.917]. Step 3 entered scenario dummy codes and participant gender as control variables, and both predicted victim gender assumptions. Participants assumed a female victim more often in response to the medical scenario compared to the retail and postal scenarios. Female participants were marginally more likely to assume the victim was female compared to male participants. However, including these variables did not eliminate the effect of labeling [B = 0.384, Wald (1) = 6.125,  $p = .013$ ; odds ratio = 1.469].

#### 4.2.3. Perceived deservingness

An ANCOVA tested the effect of offender gender and assumed victim gender (and their interaction) on victim deservingness (with label, scenario, and participant gender as covariates). The female victim was perceived as less deserving of harm ( $M_{estimated} = 1.79$ ,  $SE = 0.065$ ) than the male victim ( $M_{estimated} = 2.22$ ,  $SE = 0.128$ ),  $F(1,293) = 8.552$ ,  $p = .004$ ,  $\eta_p^2 = .028$ . However, there was no significant effect of offender gender, nor an interaction between offender and victim gender,  $F_s < 0.09$ .

#### 4.2.4. Affective and moral responses

Last, we examined the association between participants' victim gender assumptions and their affective responses to and moral judgments of the victim. Analysis of covariance (using labeling, scenario, participant gender, offender gender, and the offender/victim-gender interaction as covariates) indicated participants felt more warmly toward the victim when they assumed a female ( $M_{estimated} = 5.85$ ,  $SE = 0.06$ ) versus male victim ( $M_{estimated} = 5.49$ ,  $SE = 0.12$ ),  $F(1,293) = 7.15$ ,  $p = .008$ ,  $\eta_p^2 = .024$ . Mediation analysis (Preacher & Hayes, 2008) revealed a significant indirect effect from victim/perpetrator labeling, through assumed victim gender, to affective reactions toward that victim,  $B = 0.021$ ,  $SE = 0.021$ ,  $CI_{95} = 0.003-0.055$ . This finding is consistent with the interpretation that labeling the targets as perpetrator/victim increased participants' likelihood of assuming a female victim, which increased positive feelings towards the victim (see Fig. 1). Participants also perceived the victim as more moral when they assumed a female ( $M_{estimated} = 5.76$ ,  $SE = 0.045$ ) versus male victim ( $M_{estimated} = 5.45$ ,  $SE = 0.088$ ),  $F(1,293) = 10.106$ ,  $p = .002$ ,  $\eta_p^2 = .033$ . Mediation analysis again indicated a significant indirect effect from victim/perpetrator labeling, through assumed victim gender, to moral judgments of the victim,  $B = 0.018$ ,  $SE = 0.010$ ,  $CI_{95} = 0.004-0.045$ .

#### 4.3. Discussion

Study 1 provided preliminary support for the biased application of moral typecasting. Across three scenarios, participants were more likely to assume a harmed target was female than male, but particularly when the targets were explicitly labeled perpetrator/victim (as opposed to party A/B). This finding suggests that activating a cognitive prototype

of harm increases the likelihood of typecasting females as victims. Participants were especially likely to assume a female victim when the perpetrator was also a woman, which corresponds to extant data on workplace bullying (Workplace Bullying Institute, 2017). That is, although women are less likely to perpetrate workplace bullying, when they do, they are twice as likely to target other women than men. These patterns support the possibility that base rates of harm contribute to gendered stereotypes of victims and perpetrators. Regarding observers' gender, female participants were slightly more likely than men to assume a female victim, suggesting the hypothesized moral typecasting bias occurs more strongly among women. However, this effect was only marginally significant, so further evidence is required before we can confidently draw this conclusion. Of note, participants' assumption of a female victim shifted their affective reactions to and moral assessments of the afflicted target. Those who assumed a female (as opposed to male) victim felt more warmth toward the victim as well as perceived the victim as more moral and less deserving of harm, suggesting that when harmed individuals more (versus less) closely fit the cognitive prototype of victim, certain moral responses are amplified.

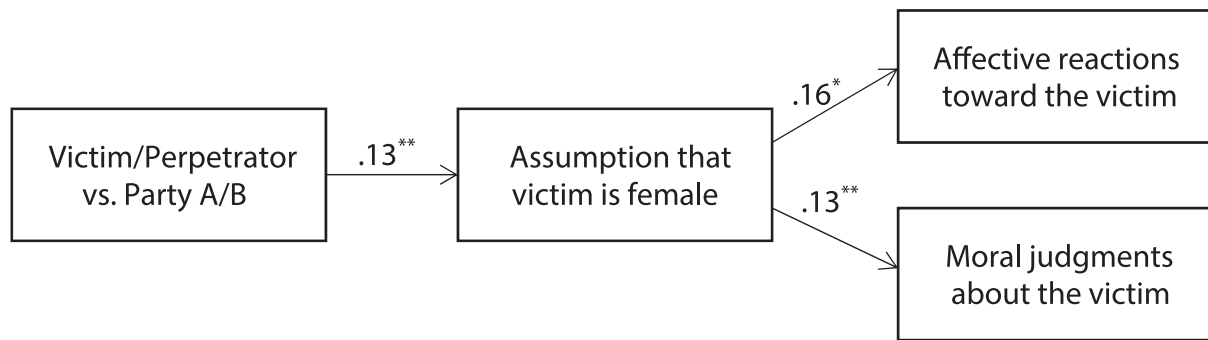
One limitation of Study 1 was that it relied on social scenarios which may have activated extraneous gender stereotypes that artificially inflated the likelihood of assuming a female victim. For example, one scenario described a surgeon bullying a surgical trainee. Knowledge of gender discrepancies in positions of power may have facilitated the assumption a female would be in a subordinate role, thereby influencing participants' expectations independently of moral typecasting. Although gender base rates may limit the generalizability of Study 1's findings, the greater assumption of a female victim in response to labeling the parties as victim/perpetrator nonetheless suggests the biased application of moral typecasting is amplified when features of harm are made salient.

### 5. Study 2

Study 2 eliminated many of the humanizing features of the stimuli (and thereby reduce the influence of extraneous gender stereotypes) to provide a more conservative test of the hypothesized gender bias in moral typecasting. Study 2 used three videos of animated shapes to represent social actors. Study 2 also extended the scope of Study 1 by examining the perpetrator side of moral typecasting to assess whether individuals would not only be more likely to typecast victims as female, but also typecast perpetrators as male. Moreover, we used a cross-cultural sample of Chinese and Norwegian participants, allowing us to evaluate whether the biased application of moral typecasting may constitute a universal feature of social cognition.

The Chinese managers and Norwegian students each evaluated a series of three brief videos depicting interactions between two triangles, ostensibly representing the interaction of a male and female coworker in an organization. We varied the specifics of the harm context to determine whether gendered assumptions are unique to one type of harm (e.g., single perpetration, where A harms B) or extend to more ambiguous forms of harm (e.g., retaliation). In response to each video, participants provided perceptions of each triangle's perpetration and victimization, and then classified each as male or female. We hypothesized that perceptions of a triangle's victimization would predict typecasting that triangle as female, whereas perceptions of perpetration would predict typecasting that triangle as male. We pre-registered these hypotheses on AsPredicted.org<sup>1</sup> (<https://aspredicted.org/tc97r.pdf>).

<sup>1</sup> The reported analyses diverged from the pre-registration because the nested nature of the data (participants viewed multiple videos and comprised two samples) required multi-level modeling.



**Fig. 1.** Study 1 – Indirect effect of victim/perpetrator labeling. Note: \*\* $p < .01$ , \* $p < .05$ ; offender gender and the labeling  $\times$  offender gender interaction are included as controls in the bootstrapping regression.

## 5.1. Method

### 5.1.1. Participants

Sample A consisted of 264 Chinese managers, enrolled in a part-time MBA program in Northern China. One hundred and four were men, (153 women, 7 missing), averaging 32.7 years of age ( $SD = 6.12$  years). Sample B consisted of 138 Norwegian university students, 110 (80%) of which were women. Thus, our total sample for our within-subject design included 402 individuals with 263 (65%) women. Due to an error in Sample B's survey, demographic data were not collected from Norwegian participants. Although we could ascertain the gender breakdown of the sample based on who accepted the study invitations, we could not link gender or age to each participant's respective data. Studies using this same Norwegian sample typically show an age range of around 20–26 years (Sjåstad & Baumeister, 2018).

### 5.1.2. Procedure

Participants evaluated three brief videos depicting the interactions of two animated triangles. Early pioneering work established that people instinctively attribute human attributes, including motivation and agency, to animated objects (Heider & Simmel, 1944). All instructions and measures were translated into each sample's native language (Chinese or Norwegian). Before viewing the videos, participants were told that the scenes represented the interaction of a male and female co-worker in an organization, and that scenes might include interactions in which one party experienced psychological, nonphysical harm as a result of the other party's behavior. We specified that the harm was not physical to avoid activating gender stereotypes about perpetrators of physical aggression. In each video, an orange triangle left its respective area (office/cubicle) to enter a common area. Next, the green triangle left its respective area to meet the orange triangle in the common area. The subsequent behavior of the triangles differed across the three videos:

**No perpetration:** In this video, the green triangle approached the orange triangle in the common area. The green triangle “looks” at the orange triangle (i.e., one of the vertices of the triangle points toward the other triangle), but no contact is depicted. The triangles faced one another (i.e., vertices pointed toward the other) before returning to their respective areas.

**Single perpetration:** In this video, the green triangle appeared to “poke” the orange triangle by making a quick move towards it (without making direct contact). The orange triangle swirls backwards in response. The two triangles then faced one another and returned to their respective work areas.

**Perpetration and Retaliation:** In this video, the green triangle again pokes the orange triangle (without direct contact), leading the orange triangle to swirl backwards. Next, the orange triangle retaliates by poking the green triangle twice (without direct contact). The green triangle swirls backwards upon each poke, after which both triangles returned to their respective work areas.

Presentation of the videos was randomized to control for order effects. Last, participants provided basic demographic data.

### 5.1.3. Dependent measures

In response to each video, participants were asked to rate the extent to which they perceived both the green and orange triangles as the victims of harm, using 7-point scales (1 = *not at all the victim*, 7 = *definitely the victim*). Participants also rated the extent to which each of the triangles was the perpetrator of harm (1 = *not at all the perpetrator*, 7 = *definitely the perpetrator*). Last, they identified the sex of the triangles by selecting from one of two options: *orange female/green male* or *orange male/green female*.

## 5.2. Results

Three-level Hierarchical Linear Models (HLM 7.01; Raudenbush, Bryk, & Congdon, 2013) were constructed to account for the nested nature of the data. Due to the within-subjects design, participants' responses to the three videos were repeated. Participants were also nested within their particular cultural sample (Chinese or Norwegian). Thus, participants' repeated responses to the videos were entered at Level 1 (along with dummy codes to represent the particular video), between-person demographics were entered at Level 2, and the cultural sample was dummy coded at level 3. Across models, all terms were fixed and not allowed to vary. Because the dependent variable (gender assignment) was dichotomous, Bernoulli specification was used, which applies a logistic link function, such that results indicate the log odds of the outcome, as opposed to the conventional raw outcome.

### 5.2.1. Effect of video on gender classification

First, the main effect of video scenario was explored by entering video scenario dummy codes at Level 1 to predict gender typecasting of the triangles. The retaliation video significantly differed from the single perpetration,  $b = -0.29$ ,  $SE = 0.14$ ,  $t(800) = 2.05$ ,  $p = .041$ , and the no overt perpetration video,  $b = -0.31$ ,  $SE = 0.14$ ,  $t(800) = 2.19$ ,  $p = .029$ , in gender classification. Participants were less likely to classify the orange triangle as female when the orange triangle retaliated, compared to when the orange triangle was only a victim or when no apparent harm took place. This pattern suggests that perpetration, even in the form of retaliation, reduced the likelihood of classification as female, supporting the hypothesis that victims are more likely to be typecast as female than agentic perpetrators.

Gender classification of triangles did not significantly differ between the single perpetration video and no overt perpetration video ( $p = .886$ ). This null effect in gender classification was unexpected, but may suggest that in response to the no overt harm video, participants inferred the green triangle's facing the orange signified a harsh glare or delivery of a cruel statement. Because instructions specified that harm was not physical, but rather psychological, it is possible participants were sensitized to detecting relational or verbal forms of harm, such as

“dirty looks” or confrontational statements. This interpretation is speculative, but if correct, may suggest that in accordance with Study 1’s findings, activating a mindset of harm amplifies the gender bias in moral typecasting, even when harm is ambiguous (e.g., a prolonged look).

### 5.2.2. Perpetration predicts gender typecasting

To test whether perceptions of victimization/ perpetration would predict gender typecasting, a series of 3-level analyses were conducted. The first model entered participants’ standardized uncentered perceptions of the green triangle’s perpetration into Level 1. Supporting our hypothesis, perceptions of the green triangle’s perpetration significantly predicted typecasting the green triangle as male,  $b = 0.29$ ,  $SE = 0.07$   $t(801) = 4.48$ ,  $p < .001$ , odds ratio = 1.34. When the video dummy codes and their interaction terms (video by perceptions of green perpetration) were also entered into Level 1 of the model, the interaction terms were not significant (all  $ps > 0.5$ ), indicating the effect of the green triangle’s perpetration on male gender assignment did not differ significantly across the three videos. To explore whether the main effect of green’s perpetration differed across the cultural samples, a study sample dummy code was entered at Level 3. Study sample did not significantly moderate the effect of green perpetration ( $p = .487$ ), indicating that Chinese managers and Norwegian students did not differ in their gender classification based on perceptions of the green triangle’s degree of perpetration.

A similar model examined perceptions of the orange triangle’s perpetration. Supporting predictions, a significant main effect of perceptions of the orange triangle’s perpetration emerged, such that the more participants perceived the orange triangle as a perpetrator, the more likely they were to classify the orange triangle as male,  $b = 0.29$ ,  $SE = 0.06$ ,  $t(801) = 4.64$ ,  $p < .001$ , odds ratio = 1.34. The strength of this association did not differ by video scenario ( $ps > 0.4$ ), nor study sample ( $p = .118$ ).

### 5.2.3. Victimization predicts gender typecasting

Next, perceptions of the green triangle’s victimization were examined. Supporting predictions, the more participants perceived the green triangle as a victim, the greater their likelihood of typecasting the green triangle as female,  $b = 0.40$ ,  $SE = 0.06$ ,  $t(801) = 6.28$ ,  $p < .001$ , odds ratio = 1.50. This main effect was not qualified by the study sample ( $p = .679$ ), indicating that the pattern was similar among Chinese and Norwegian participants. The effect of green triangle’s victimization significantly differed across the particular videos, however, such that effects significantly differed between the *no perpetration* and *single perpetration* scenarios,  $b = 0.55$ ,  $SE = 0.19$ ,  $t(797) = 2.91$ ,  $p = .001$ . Perceptions of green’s victimization was a stronger predictor of classifying the green triangle as female in the single perpetration video,  $b = 0.76$ ,  $SE = 0.15$ ,  $t(797) = 5.11$ ,  $p < .001$ , odds ratio = 2.13, than in the video depicting no overt perpetration,  $b = 0.21$ ,  $SE = 0.13$ ,  $t(797) = 1.63$ ,  $p = .104$ , odds ratio = 1.23. In the single perpetration video, it was quite clear that the green triangle was the aggressor, suggesting that if participants happened to perceive the green triangle as a victim in this scenario, they were especially likely to typecast the green triangle as female.

Turning to perceptions of the orange triangle, results revealed a significant main effect of victim perceptions, such that the more participants perceived the orange triangle as a victim, the greater their likelihood of typecasting the orange triangle as female,  $b = 0.27$ ,  $SE = 0.06$ ,  $t(801) = 4.15$ ,  $p < .001$ , odds ratio = 1.31. The strength of this association did not differ across the videos ( $ps > 0.3$ ), nor across study samples ( $p = .535$ ).

### 5.2.4. Participant gender and typecasting

A final set of analyses examined the effect of participant gender on typecasting. Because demographic data were not collected for the Norwegian sample, this analysis was performed solely on the Chinese

managerial sample, and thus only required a two-level model. A participant gender dummy code was entered into Level 2 to examine the main effect of participant gender on typecasting. A significant effect emerged,  $b = 0.45$ ,  $SE = 0.18$ ,  $t(255) = 2.49$ ,  $p = .013$ , odds ratio = 1.57, indicating that female participants were more likely to classify the green triangle as male and the orange triangle as female than were male participants. Because the green triangle was the sole perpetrator in one video and the initial aggressor in another, this finding may suggest women were more likely than men to assume male perpetration and female victimization.

## 5.3. Discussion

Study 2 supported the predicted biased application of moral typecasting using stimuli devoid of human attributes. Even when judging animated shapes, the more participants perceived a triangle as a victim in a social interaction involving harm, the more likely they were to classify that triangle as female. Likewise, the more participants perceived a triangle as a perpetrator, they more likely they were to typecast that triangle as male. This pattern of typecasting was found across all three videos, including the retaliation scenario, suggesting biased typecasting can manifest even when victimization is ambiguous. Study 2’s stimuli reduced the possibility the emergent pattern stemmed primarily from gender stereotypes because participants attributed these human characteristics to animated shapes. Moreover, there were no significant differences between Chinese managers’ and Norwegian students’ responses, suggesting the biased application of moral typecasting may be a universal feature of human moral cognition.

Exploration of participant gender showed that female participants were more likely than male participants to classify the orange triangle as female and the green as male. Because the green triangle was the sole perpetrator in one video and the initial aggressor in the other, this finding suggests that women may be more likely than men to reflexively assign males to the role of perpetrator and females to the role of victim. However, because the significant effect of participant gender was not unique to the single perpetration video (where victimization and perpetration were clear) but rather occurred across all three videos, we are cautious about interpreting this as a reliable pattern. Study 2’s design was limited in that the interaction was described as between opposite-sex colleagues and thus, did not allow us to explore whether biased moral typecasting emerges when evaluating male-on-male or female-on-female harm.

## 6. Study 3

Study 2’s findings suggested that biased application of moral typecasting persists even when victimization is ambiguous (e.g., in the retaliation video). Study 3 sought to further test this possibility with a scenario depicting workplace conflict between two opposite-sex colleagues. In this scenario, both employees experienced negative outcomes, and thus, assigning actors to the victim role was not straightforward.

A second goal of Study 3 was to examine a possible boundary condition of the biased application of moral typecasting by increasing the patency of one of the targets, which should also decrease their perceived agency. We did so by manipulating whether an overly friendly co-worker was described as having a neurological condition that impaired his/her ability to interpret social cues. Prior work has demonstrated that persons with mental handicaps are attributed lower agency and heightened patency (Gray & Wegner, 2009). If gender is used as only one proxy of individuals’ relative agency or patency, then perhaps cues suggestive of diminished agency and augmented patency (such as neurological impairment) might overpower the biasing effects of target gender on moral typecasting. Accordingly, we predicted that the female target would be perceived as more of a victim than the male in the control condition, but not in the neurological impairment

condition. Thus, Study 3 used a 2 (target gender) by 2 (neurological impairment) design to examine moral typecasting when victimization is ambiguous.

## 6.1. Method

### 6.1.1. Participants

Two hundred and nineteen Chinese managers ( $M_{\text{age}} = 31.8$  years, range: 20–56) enrolled in a part-time MBA program in Northern China were recruited to complete an online survey. Of these managers, 87 (39.7%) were male and two did not report gender.

### 6.1.2. Procedure.

Participants read a short scenario depicting the workplace interactions of two opposite sex co-workers. Within the scenario, the first co-worker was described as very friendly, often giving hugs and compliments to others. The scenario detailed how the friendly coworker often stopped by their co-worker's cubicle to chat, often sent emails unrelated to work, frequently stared at the co-worker, and once gave an unwelcomed hug when their co-worker came into work soaking wet on a rainy day. The co-worker who was the recipient of this attention became distressed and eventually emailed the company's HR documenting all the ways the friendly employee made coming into work uncomfortable for him/her. A few days later, the overly friendly employee was let go.

Across conditions, the sex of the two colleagues were manipulated, such that either a female employee behaved overly friendly towards a male co-worker, leading to her being fired or a male employee behaved in an overly friendly manner towards a female co-worker, leading to his being fired. We also manipulated the neurological status of the friendly employee, such that (s)he was described as having a neurological impairment that harmed his/her ability to read social cues, or in the control condition, there was no mention of a neurological impairment.

### 6.1.3. Dependent measures

In response to the vignette, participants indicated on two 7-point scales the degree to which they perceived the fired person as a victim and the accuser as a victim (1 = *not at all a victim*, 7 = *definitely a victim*).

## 6.2. Results

### 6.2.1. Complainant's victimization

Participants reported their perception of both targets' victimization and their responses to these two items are therefore interrelated. That is, because individuals typically typecast one target as a victim and one as a perpetrator within harm contexts (Gray & Wegner, 2009), victimization ratings for one target may have been confounded by ratings of the other target. To account for this within-subject variance on these two measures, we conducted ANCOVAs that accounted for the rating of the other target's victimization as a covariate. The first ANCOVA examined perceptions of the complainant's victimization by gender condition, neurological condition, and participant gender, while controlling for perceptions of the fired employee's victimization. The predicted 2-way gender by neurological condition was significant,  $F(1,208) = 7.35$ ,  $p = .007$ . In the control condition, the female complainant was perceived as more of a victim ( $M = 4.12$ ,  $SE = 0.20$ ) than the male complainant ( $M = 3.03$ ,  $SE = 0.19$ ),  $p < .001$ . In the neurological impairment condition, there was no difference in perceptions of the complainant's victimization based on gender ( $M_{\text{male}} = 3.80$ ,  $SE = 0.22$ ;  $M_{\text{female}} = 3.77$ ,  $SE = 0.21$ ),  $p = .923$ . This two-way interaction was not moderated by participant gender,  $F(1,208) = 0.61$ ,  $p = .435$ , suggesting male and female participants both showed a gender bias in their assessments of the complainant's victimization, but only in the no-impairment condition. This finding supported predictions, demonstrating that a female employee making a complaint about

being distressed by another co-worker's behavior was perceived as more of a victim than a male employee making the same complaint, but this gender bias was not found when the accused suffered from a neurological impairment.

### 6.2.2. Fired Employee's victimization

A second ANCOVA examined perceptions of the fired employee's victimization, treating perceptions of the complainant's victimization as a covariate. A significant 2-way participant gender by gender condition emerged,  $F(1,208) = 4.71$ ,  $p = .031$ , such that male participants viewed the fired employee as marginally more of a victim as a female ( $M = 4.61$ ,  $SE = 0.23$ ) than as a male ( $M = 4.01$ ,  $SE = 0.23$ )  $p = .071$ , whereas female participants did not significantly differ in their ratings of victimization as a function of the fired employee's gender ( $M_{\text{female}} = 3.54$ ,  $SE = 0.18$ ;  $M_{\text{male}} = 3.85$ ,  $SE = 0.18$ ),  $p = .227$ . This pattern suggests that male, but not female participants showed a gender bias in their typecasting of the fired employee, perceiving the female employee who was fired for her overly friendly behavior as more of a victim than a male employee who engaged in identically friendly behavior. A marginally significant participant gender by neurological condition also emerged  $F(1,208) = 3.18$ ,  $p = .076$ , such that male participants perceived the fired employee as more of a victim when that employee had a neurological impairment ( $M = 4.53$ ,  $SE = 0.24$ ), than did female participants ( $M = 3.56$ ,  $SE = 0.19$ ),  $p = .001$ . In the control condition, male ( $M = 4.09$ ,  $SE = 0.21$ ) and female participants ( $M = 3.84$ ,  $SE = 0.18$ ) did not significantly differ in their assessments of the fired person's victimization,  $p = .373$ . This pattern suggests the neurological impairment increased perceptions of the fired employee's patency (and thus victimization), but only among male participants.

### 6.2.3. Comparison of employees' victimization

Given that victimization in the scenario was ambiguous, and moral typecasting assumes that only one target is typically recognized as a victim, we conducted a repeated-measure mixed ANOVA to compare victimization perceptions across the two targets. Victimization ratings for the fired and complaining employee were treated as a within-subject (or repeated) factor while gender condition, neurological condition, and participant gender were treated as between-subject factors. This analysis revealed a significant interaction between target gender and the comparison between the fired employee's and complainant's victimization,  $F(1,209) = 5.61$ ,  $p = .019$ . When the fired employee was female, she was perceived as more of a victim ( $M = 4.09$ ,  $SE = 0.15$ ) than the male complainant ( $M = 3.41$ ,  $SE = 0.15$ ),  $p = .002$ . However, when the fired employee was male, he was not perceived as more of a victim ( $M = 3.92$ ,  $SE = 0.14$ ) than the female accuser ( $M = 3.95$ ,  $SE = 0.15$ ),  $p = .897$ . Thus, a female employee who was terminated for ambiguously friendly behavior was perceived as more of a victim than her accuser, but no such difference was found in the reverse scenario, when a male employee was fired for equally ambiguous behavior.

However, this interaction was qualified by a significant interaction between target gender and participant gender on the comparison between the fired employee's and complainant's victimization,  $F(1,209) = 13.95$ ,  $p < .001$ . Male participants perceived the fired female employee as significantly more of a victim ( $M = 4.63$ ,  $SE = 0.23$ ) than the male complainant ( $M = 3.14$ ,  $SE = 0.23$ ),  $p < .001$ . When a male employee was fired for the same ambiguous behavior, male participants did not perceive him as more of a victim ( $M = 3.99$ ,  $SE = 0.22$ ) than the female complainant ( $M = 4.30$ ,  $SE = 0.23$ ),  $p = .334$ . Female participants viewed both the complainant and fired employee similarly, regardless of the gender manipulation ( $ps > 0.32$ ).

The predicted interaction between neurological condition, gender condition, and the employees' victimization ratings was marginally significant,  $F(1,209) = 2.96$ ,  $p = .087$ . Probing this interaction revealed that in the neurologically healthy condition, the female complainant was perceived as more of a victim ( $M = 4.12$ ,  $SE = 0.20$ ) than the male complainant ( $M = 3.03$ ,  $SE = 0.19$ ),  $p < .001$ . However, in



the neurological impairment condition, the female complainant was no longer perceived as more of a victim ( $M = 3.78$ ,  $SE = 0.21$ ) than the male complainant ( $M = 3.80$ ,  $SE = 0.22$ ),  $p = .948$ .

### 6.3. Discussion

Study 3's findings suggest that individuals show a biased application of moral typecasting even under conditions of ambiguity where it is not obvious who is the victim. A woman was perceived as more of a victim than a man making the same complaints of distressing, unwanted attention from a coworker. This result supports our argument that it is more cognitively challenging to typecast a man than a woman into the victim role. This tendency can lead decision makers to stray from impartiality when resolving workplace disputes because complaints from male employees may be taken less seriously. Unlike our previous studies, the biased application of moral typecasting was most pronounced among male participants, who perceived the fired employee as more of a victim when female than male. Indeed, male participants also perceived the fired female employee as more of a victim than the male employee making complaints about her.

Of note, the findings of Study 3 suggest a boundary condition to the biased application of moral typecasting. When the fired employee suffered from a neurological impairment, the gender bias in moral typecasting disappeared, and both male and female employees making complaints of harassing behaviors were assigned equal victim status. We infer from this that gender is only one factor that biases the application of moral typecasting due to its association with agency and patiency. However, in the presence of additional and perhaps more diagnostic cues of relative patiency or agency, such as a neurological impairment, evaluators may weigh these cues more heavily than targets' gender. One limitation of both Studies 2 and 3 was that both relied upon scenarios depicting cross-sex interactions. We addressed this limitation in Study 4.

## 7. Study 4

Study 4 manipulated both the genders of a potentially harmed target and a perpetrator to examine the downstream consequences of the biased application of moral typecasting. If women more strongly match the cognitive prototype of victim than men, female targets should be expected to experience more pain than male targets, consistent with the patient role in moral typecasting (Gray & Wegner, 2009). Study 4 directly tested this prediction by measuring assumptions about a victim's pain. Study 4 also explored reactions to male versus female perpetrators. The cognitive link between female and victimization demonstrated in our previous studies implies that even when women are cast as perpetrators, they should be assumed to experience more pain (e.g., upon being accused of harassment) compared to when men are cast in this role. In other words, even in the perpetrator role, female targets will still be attributed qualities of victims, making it more challenging for evaluators to detect and respond punitively to female (versus male) perpetration. Specifically, we predicted that female perpetrators would elicit less punitive responses than male perpetrators. Last, we predicted that perpetrators who harmed female victims should evoke more punitive responses and lower inclinations for forgiveness than those who harmed male victims.

### 7.1. Method

#### 7.1.1. Participants

Two hundred and fifty American residents were recruited to participate in an online study through Amazon's MTurk ( $M_{age} = 34.49$ , 34% female). Participants were first provided with definitions of workplace, sexual, and quid pro quo harassment, and asked whether they had ever been formally or informally accused of one of these. Those who responded "yes" were directed to a survey about their experience being

accused (findings reported in a separate article); those who responded "no" were directed to the study reported here. Our final sample consisted of 214 individuals who indicated 'no' ( $M_{age} = 34.93$ , 38% female).

#### 7.1.2. Procedure

Participants were randomly assigned to view one of four versions of a dyadic workplace interaction that manipulated both commenter and target gender. Participants were presented with the following scenario: "A and B are from the same company and they have worked together for more than two years. Both are managers. Occasionally, after work they go out for drinks with several colleagues. One day, A and B attend a professional conference. At the conference, during the lunch break, A and B are in the cafeteria line together. B drops a fork and bends over to get it. As (s)he straightens up, A says to her/him, 'You must get a lot of practice doing that.'" The genders of A and B were experimentally manipulated using names (e.g., Jason, Diane) and pictures.

We used this particular joke for two reasons. First, there is no way of objectively determining whether the joke is offensive. Second, if perceived as offensive, the joke is potentially hurtful to both men and women because it implies bodily penetration. Additionally, when the comment is directed towards a man, it can be interpreted as emasculating because it implies anal penetration, presumably by another man. Thus, it would be reasonable to expect that if the comment were perceived as offensive, it would be so for both men and women.

#### 7.1.3. Dependent measures

After reading the scenario, participants completed the following measures:

**7.1.3.1. Recipient pain.** To measure perceptions of the recipient's (B) pain in response to the comment, participants responded to the question 'How much pain do you think B felt as a result of how A acted towards him/her?' using a slider (0 = no pain and 100 = extreme pain).

**7.1.3.2. Commenter pain.** To measure perceptions of the commenter's (A) pain, participants responded to the question 'Assuming B did accuse A of harassment, how much pain do you think A would feel as a result of B's accusation?' using a slider (0 = no pain and 100 = extreme pain).

**7.1.3.3. Punishment.** Participants indicated their desire to punish the commenter (A) on five items using a 5-point scale (1 = definitely not, 2 = probably not, 3 = neutral/unsure, 4 = probably, 5 = definitely). "Assuming B went to A's supervisor and accused him/her of creating a hostile work environment for (wo)men because of his/her behavior in his/her presence, indicate how much you believe A's supervisor should take each of the following steps: 1) A should be punished in some way for his/her behavior; 2) A's supervisor should pursue a thorough investigation of A's behavior at work to see if (s)he has done this to other (wo)men; 3) A's supervisor should suspend him/her without pay while conducting an investigation about B's complaint; 4) A's supervisor should require A to seek counseling for his/her behavior; 5) A's supervisor should fire A for his/her behavior." Responses were combined to form a punishment composite ( $\alpha = 0.89$ ).

**7.1.3.4. Forgiveness.** We measured participants' likelihood of forgiving the commenter (A) with four statements and the same 5-point response scale: 1) I would let go of any of the negative feelings I have towards A; 2) I would forgive A for any hurt and pain (s)he may have caused; 3) I would let go of any resentment that I felt towards A; 4) I would never forgive A for what (s)he said to B (reverse coded). Responses were combined to form a forgiveness composite ( $\alpha = 0.91$ ).

**7.1.3.5. Workplace advancement.** Two items examined participants' desired workplace advancement for the commenter: 1) Would you want to work with [Commenter]? and 2) Would you recommend

[Commenter] for a management position? using the same 5-point scale ( $\alpha = 0.89$ ).

## 7.2. Results

A series of 2 (recipient gender)  $\times$  2 (commenter gender) ANOVAs examined participants' responses to the commenter and recipient of a workplace joke. A secondary set of ANOVAs examined moderation by participant gender. Although main effects of participant gender emerged, such that female (versus male) participants perceived greater victim pain, desired harsher punishments, and were less willing to forgive or integrate offenders, there were no significant interactions between participant gender, commenter gender, or recipient gender. Thus, we report only the 2 (recipient gender)  $\times$  2 (commenter gender) ANOVAs below.

### 7.2.1. Victim's pain

Participants perceived female recipients to experience more pain than male recipients ( $M_{\text{female}} = 31.09$ ,  $SE_{\text{female}} = 2.35$ ,  $M_{\text{male}} = 22.14$ ,  $SE_{\text{male}} = 2.35$ ;  $F(1, 210) = 7.28$ ,  $p = .008$ ,  $d = 0.37$ ). Although the main effect of commenter gender was not statistically significant,  $F(1, 210) = 2.73$ ,  $p = .10$ , it was trending in the predicted direction, such that recipients were expected to experience more pain as a response to male (versus female) commenters. There was no significant recipient by commenter gender interaction,  $F(1, 210) = 0.17$ ,  $p = .69$ .

### 7.2.2. Perpetrator's pain

Participants perceived female commenters to experience more pain than male commenters ( $M_{\text{female}} = 58.98$ ,  $SE_{\text{female}} = 2.71$ ,  $M_{\text{male}} = 49.09$ ,  $SE_{\text{male}} = 2.73$ ,  $F(1, 210) = 6.60$ ,  $p = .011$ ,  $d = 0.35$ ) upon accusation of harassment. There was no effect of victim gender,  $F(1, 210) = 0.01$ ,  $p = .94$ , and no victim by perpetrator gender interaction,  $F(1, 210) = 0.00$ ,  $p = 1.0$ .

### 7.2.3. Punishment

Participants desired harsher punishments for perpetrators who targeted females ( $M_{\text{femaleV}} = 2.74$ ,  $SE_{\text{femaleV}} = 0.09$ ) than those who targeted males, ( $M_{\text{maleV}} = 2.49$ ,  $SE_{\text{maleV}} = 0.09$ ,  $F(1, 210) = 3.73$ ,  $p = .055$ ,  $d = 0.27$ ). A main effect of perpetrator gender revealed that participants were more willing to punish male vs. female perpetrators ( $M_{\text{maleP}} = 2.77$ ,  $SE_{\text{maleP}} = .09$ ,  $M_{\text{femaleP}} = 2.47$ ,  $SE_{\text{femaleP}} = .09$ ,  $F(1, 210) = 5.35$ ,  $p = .022$ ,  $d = 0.32$ ). There was no victim by perpetrator gender interaction,  $F(1, 210) = 0.02$ ,  $p = .89$ .

### 7.2.4. Forgiveness

Participants were more willing to forgive a female than male perpetrator ( $M_{\text{femaleP}} = 3.80$ ,  $SE_{\text{femaleP}} = .09$ ,  $M_{\text{maleP}} = 3.48$ ,  $SE_{\text{maleP}} = .09$ ,  $F(1, 210) = 7.05$ ,  $p = .009$ ,  $d = 0.34$ ). There was no main effect of victim gender,  $F(1, 210) = 1.32$ ,  $p = .25$ , and no victim by perpetrator gender interaction,  $F(1, 210) = 0.85$ ,  $p = .36$ .

### 7.2.5. Workplace advancement

A significant effect of perpetrator gender emerged,  $F(1, 210) = 10.84$ ,  $p = .001$ , such that participants were more willing to work with and advocate for the promotion to management of a female than male perpetrator ( $M_{\text{maleP}} = 2.39$ ,  $SE_{\text{maleP}} = .14$ ,  $M_{\text{femaleP}} = 2.83$ ,  $SE_{\text{femaleP}} = .13$ ). There was no main effect of victim gender,  $F(1, 210) = 1.31$ ,  $p = .25$ , nor a victim by commenter gender interaction,  $F(1, 210) = 0.09$ ,  $p = .76$ .

### 7.2.6. Mediation analysis

To test the hypothesis that people more strongly desire to punish perpetrators who target female (vs. male) victims because they perceive female victims' pain to be greater than that of male victims, we tested a mediation model using Model 4 in the Hayes Macro for SPSS (Hayes, 2017) with victim gender as the IV, perceived victim pain as the

mediator, and punishment as the DV. We controlled for participant gender in the model.

The victim gender-victim pain pathway was significant, indicating that participants believed a female victim experienced significantly more pain from the joke than a male victim,  $t(210) = 2.96$ ,  $p = .004$ ,  $CI_{95} = [3.26, 16.31]$ . The pathway from victim pain to perpetrator punishment was also significant such that higher perceived victim pain predicted harsher perpetrator punishment,  $t(210) = 10.92$ ,  $p < .001$ ,  $CI_{95} = [0.019, 0.027]$ . Last, the direct effect of victim gender on perpetrator punishment was no longer significant,  $t(210) = .53$ ,  $p = .60$ ,  $CI_{95} = [-0.151, 0.261]$ , but the indirect effect was,  $B = 0.227$ ,  $CI_{95} = [0.067, 0.395]$ , suggesting people perceived female victims to experience greater pain, leading to stronger punitive responses toward their perpetrators. Participant gender was a significant covariate such that female participants assumed greater victim pain,  $t(210) = 2.34$ ,  $p = .02$ ,  $CI_{95} = [1.27, 14.72]$ , but participant gender was not associated with punishment desires,  $t(210) = 1.53$ ,  $p = .127$ ,  $CI_{95} = [-0.047, 0.374]$ .

## 7.3. Discussion

Study 4 revealed some of the practical consequences of the biased application of moral typecasting. Participants perceived females to experience more pain than males in response to the same potentially offensive joke. The dyadic template framework contends that typecasting a target as a victim makes it unlikely that they will be typecast as an agent. Our finding that female (versus male) perpetrators were also expected to experience greater pain (a component of moral patiency) is consistent with this principle.

Participants were less willing to punish and more willing to forgive female than male perpetrators, despite their committing the same offense. They were more willing to work with and nominate a woman for a position of management than a man, despite both making the same potentially offensive joke. Participants also desired harsher punishments, including investigations and terminations, for those who targeted female (versus male) victims. Thus, it appears that those alleged to harm females are perceived as especially pernicious and punished more severely for their actions. Mediation analysis revealed these greater punitive desires stemmed in part from an assumption that female victims experienced more pain than male victims, suggesting a mechanism by which this asymmetry occurs. In sum, Study 4 revealed third parties experience stronger moral responses when actors' genders more closely fit the cognitive prototype of intentional perpetrator and suffering victim.

## 8. Study 5

Thus far, our studies have focused exclusively on dyadic interactions involving one party experiencing harm. It remains unclear whether biases in moral typecasting will persist when more than one target is harmed. Study 5 expanded the scope of our investigation by examining whether observers exhibit a gender bias in moral typecasting when evaluating situations depicting harm to either women or men as a group. To provide a conservative test of the biased application of moral typecasting, Study 5 sought to rule out the influence of gender stereotypes as the primary driver of our effects by asking participants to respond to a situation where stereotypes could facilitate men's categorization into the victim role: job loss. Throughout history, men have been perceived as the primary household breadwinners, suggesting that losing a job may more strongly impair men's sense of self-worth than women's, thereby resulting in men's experiencing more pain. Empirical data provide support for this expectation as job instability more strongly predicts depression and gender threat among men than among women (Michniewicz, Vandello, & Bosson, 2014; Wang, Lesage, Schmitz, & Drapeau, 2008).

## 8.1. Method

### 8.1.1. Participants

A market research firm recruited 423 Canadian participants for an online study about a managerial decision. Twenty individuals failed to identify correctly the gender of the fired employees from the scenario and were thus eliminated from analyses. Our final sample was 403 individuals (51.4% female;  $M_{\text{age}} = 45.3$  years). The majority (73.5%) were currently employed in organizational leadership roles, with an average of 22.9 years of work experience.

### 8.1.2. Procedure

Participants evaluated a scenario in which the senior management team at a manufacturing company decided to protect profits by laying off 9 employees. Across conditions, we experimentally manipulated the gender of the laid-off employees, such that they were either all men or all women. Within both conditions, employee tenure, education, and outcomes (e.g., severance; re-employment), as well as the management team's motivations were held constant. The full scenario is available in the appendix. After reading the scenario, participants completed the following dependent measures:

**8.1.2.1. Victimhood.** *The (wo)men who were laid off are NOT victims (1 = Strongly Disagree; 7 = Strongly Agree).*

**8.1.2.2. Victim pain.** *How much pain, whether it is psychological, emotional, or physical, do you think the (wo)men who were laid off experienced as a group? (0 = No Pain; 100 = Extreme Pain).*

**8.1.2.3. Harm inflicted.** *How much harm do you think the management team inflicted on the group of (wo)men they laid off? (0 = No Harm; 100 = Extreme Harm).*

**8.1.2.4. Fairness.** *I believe the (wo)men have been treated fairly (1 = Strongly Disagree; 7 = Strongly Agree).*

**8.1.2.5. Managers' morality.** *How moral do you think the senior management team at Jarvis Manufacturing is? (1 = Extremely Immoral; 7 = Extremely Moral).*

**8.1.2.6. Management team's gender composition.** *Do you think the management team who decided to make the layoffs is: (All men, All women, A mix of men and women).* Order of response options were randomized.

## 8.2. Results

### 8.2.1. Victimhood

An independent-samples *t*-test revealed that perceptions of victimhood did not significantly differ across gender conditions ( $M_{\text{women}} = 3.79$ ,  $SD_{\text{women}} = 1.91$ ,  $M_{\text{men}} = 3.88$ ,  $SD = 1.92$ ,  $t(401) = 0.46$ ,  $p = .65$ ), contrary to predictions. Because the victimhood measure was worded as a negation ("not victims"), these means are still in the predicted direction.

### 8.2.2. Victim pain

Perceptions of the employees' pain (measured using a 0–100 slider) displayed significant negative skew (Skew =  $-1.2$ , Shapiro-Wilk test of normality =  $0.91$ ,  $p < .001$ ). We therefore analyzed this variable using non-parametric tests of median and mean condition differences. Supporting predictions, median perceived pain was significantly higher in the fired female employee condition than in the male condition (Median<sub>female</sub> =  $80.0$ , Median<sub>male</sub> =  $74.0$ ,  $p = .015$ ), and mean perceived pain followed the same pattern (Mean Rank<sub>women</sub> =  $212.43$ , Mean Rank<sub>men</sub> =  $190.77$ , Mann-Whitney  $U = 18094.00$ ,  $p = .062$ ).

### 8.2.3. Harm inflicted

Perceptions of the harm inflicted by the management team also displayed significant negative skew (Skew =  $-0.67$ , Shapiro-Wilk test of normality =  $0.95$ ,  $p < .001$ ), so we again employed nonparametric tests. Supporting predictions, participants assumed the managers inflicted more harm on the fired female than fired male employees (Median<sub>female</sub> =  $71.0$ , Median<sub>male</sub> =  $69.5$ ,  $p = .051$ ), and mean perceived harm followed the same pattern (Mean Rank<sub>women</sub> =  $209.31$ , Mean Rank<sub>men</sub> =  $194.13$ , Mann-Whitney  $U = 18745.50$ ,  $p = .192$ ).

### 8.2.4. Fairness

An independent-samples *t*-test revealed participants perceived the laid off female employees were treated marginally less fairly than the male employees ( $M_{\text{women}} = 3.60$ ,  $SD_{\text{women}} = 1.75$ ,  $M_{\text{men}} = 3.91$ ,  $SD = 1.71$ ,  $t(401) = 1.76$ ,  $p = .079$ ). In other words, participants thought that it was less fair for the managers to fire women to maximize profits than to fire men.

### 8.2.5. Managers' morality

An independent-samples *t*-test revealed that participants perceived the management team as significantly less moral when they fired women than when they fired men [ $M_{\text{women}} = 3.71$ ,  $SD_{\text{women}} = 1.42$ ,  $M_{\text{men}} = 4.07$ ,  $SD = 1.43$ ,  $t(401) = 2.49$ ,  $p = .013$ ]. Of all of our dependent measures, this was the only variable to significantly differ across conditions by participant gender,  $F(1, 399) = 12.37$ ,  $p = .013$ . A 2 (fired employee gender) by 2 (participant gender) ANOVA revealed female participants perceived the managerial team to be less moral when they laid off women ( $M = 3.41$ ,  $SE = 0.14$ ) than when they laid off men ( $M = 4.10$ ,  $SE = 0.14$ ). However, male participants did not perceive the managerial team differently depending on whether they fired women ( $M = 4.04$ ,  $SE = 0.14$ ) or men ( $M = 4.03$ ,  $SE = 0.15$ ).

### 8.2.6. Management team gender composition

Supporting the well-documented masculine construal of leadership (Koenig, Eagly, Mitchell, & Ristikari, 2011), very few participants assumed the management team was all women. The number of participants who did assume an all-female management team was virtually identical between conditions ( $N = 7$  in the fired female condition;  $N = 8$  in the fired male condition), so we confined our analyses to only those who assumed the management team was either all men or mixed gender. A Chi-Square test revealed participants were slightly, but not significantly, more likely to assume the management team was all men when female employees were fired (49.5%) compared to when male employees were fired (41.9%,  $X^2_{(1)} = 2.24$ ,  $p = .135$ ).

### 8.2.7. Perceived morality by management gender assumption

To test the hypothesis that an all-male management team would be perceived as less moral than a mixed gender management team, a 2 (condition: female vs. male layoffs) by 2 (perceived management gender: all men vs. mixed gender) between-subjects ANOVA compared perceptions of the management team's morality. The significant main effects of the fired employees' gender,  $F(1, 384) = 5.53$ ,  $p = .019$ , and assumed management gender,  $F(1, 384) = 26.62$ ,  $p < .001$ , were qualified by a significant two-way interaction,  $F(1, 384) = 6.36$ ,  $p = .012$  (see Fig. 2). Contrast analyses revealed participants who assumed an all-male management team perceived the team as significantly less moral when they laid off women ( $M = 3.24$ ,  $SE = 0.14$ ) than when they laid off men, [ $M = 3.88$ ,  $SE = 0.15$ ;  $t(176) = 3.19$ ,  $p = .002$ ]. Participants who assumed a mixed-gender management team, however, did not evaluate the team differently depending on whether they fired women ( $M = 4.25$ ,  $SE = 0.13$ ) or men [ $M = 4.24$ ,  $SE = 0.13$ ;  $t(208) = -0.13$ ,  $p = .90$ ; see Fig. 2].

## 8.3. Discussion

Study 5 showed that the gender bias in moral typecasting found in

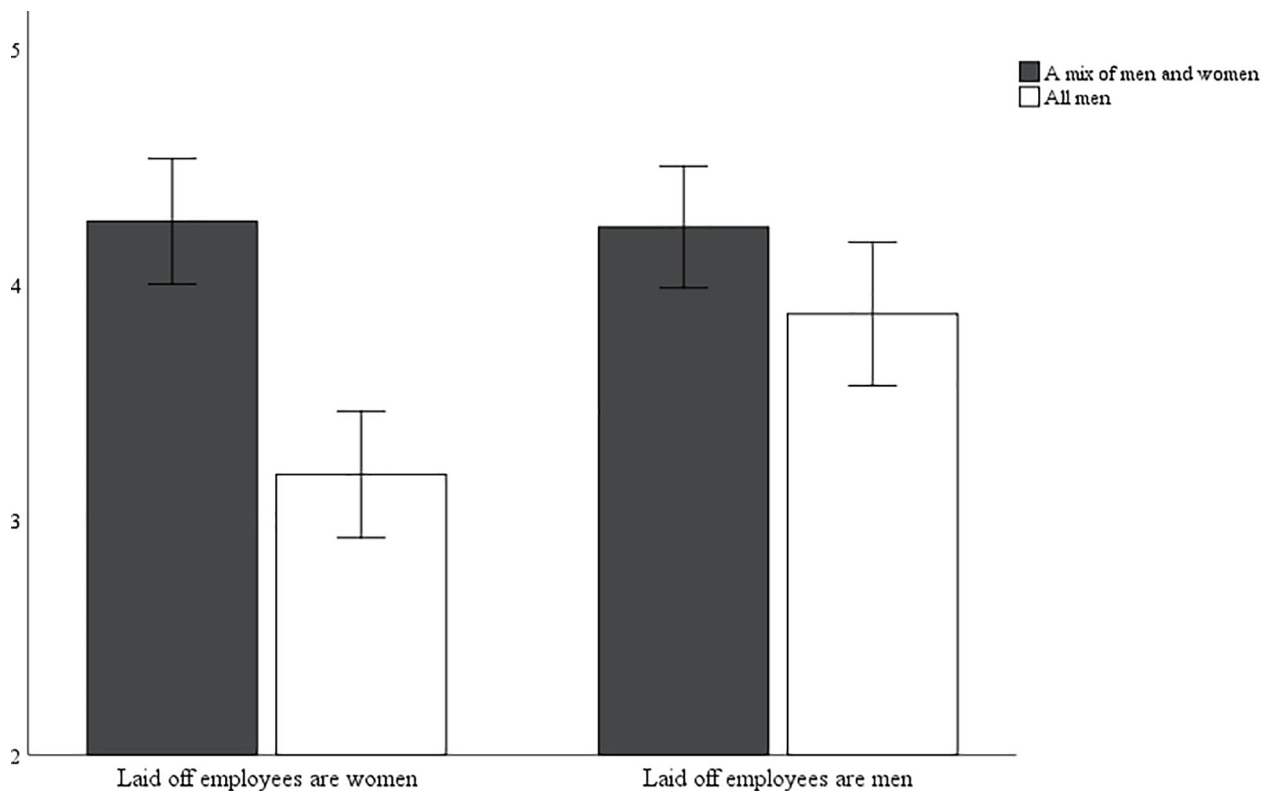


Fig. 2. Perceptions of Managerial Team's Morality By Gender of Fired Employees and Assumed Managerial Gender (Study 5).

our first four studies is not restricted to dyadic interactions involving single individuals, but aggregates to harm suffered by groups of individuals. Despite the gender stereotype of the male breadwinner and empirical research documenting the harmful consequences of job instability on men's mental health and self-image (Michniewicz et al., 2014; Wang et al., 2008), perceivers nevertheless attributed greater suffering to laid off women than to laid off men. Moreover, the perception of greater female harm also influenced evaluations of ostensible perpetrators: the management team. Organizational leaders cannot avoid making tradeoffs that often result in negative externalities, such as lost jobs. The findings of Study 5 suggest observers may judge managers as more immoral and unfair when such decisions cause harm to female than to male employees. However, female participants were especially likely to view the managerial team as immoral when they laid off women, suggesting again that women may show stronger asymmetries in moral typecasting than men.

Study 5's results suggest the gender makeup of management teams also shapes evaluations of these decisions. Supporting our predictions, participants judged an all-male team, but not a mixed-gender team, as less moral when they fired female compared to male employees. This pattern suggests that all-male management teams more easily conform to the perpetrator prototype than teams including women, and their decisions are viewed as especially pernicious and immoral if harm befalls women. This pattern should be interpreted with caution because it was correlational and not manipulated (i.e., participants assumed the managerial team's gender composition), but it nonetheless has important implications for organizations. For example, perhaps the presence of female leaders making decisions that harm women can insulate managerial teams from hostile criticism or judgments of immorality. Study 5 was limited by its reliance upon a scenario depicting layoffs from an occupation that may be considered male-dominated (i.e., ski apparel manufacturing). Thus, it is possible that perceptions of gender discrimination may have contributed to the relatively stronger responses to the firing of a group of women (versus men). We sought to rule out this explanation in Study 6.

## 9. Study 6

Study 6 employed a similar design to Study 5 whereby participants evaluated a scenario in which a managerial team fired either all female or all male employees. Unlike Study 5, however, Study 6 manipulated the plausibility of inferring gender discrimination by altering the occupation from which the employees were fired, such that they were fired either from a predominately female career (nursing) or predominately male career (chemical manufacturing; Bureau of Labor Statistics, 2018). We reasoned that the firing of a group of employees in an occupation where they are underrepresented can be used by observers as rationale for inferring discrimination because it adversely impacts presumably disadvantaged individuals from entering that field.

If there is no typecasting bias and inferences of discrimination as a result of adverse impact can fully explain the results of Study 5, then moral responses should be amplified when either women or men are fired from occupations dominated by the other gender. That is, the plausibility of gender discrimination should amplify the likelihood that both fired men and women are labeled as victims and presumed to suffer. If there is a gender bias in moral typecasting, however, then responses should be heightened when women, but not men, are fired, irrespective of plausible discrimination. A third possibility is that gender discrimination amplifies the effects of moral typecasting by strengthening the cognitive link between women and victimhood, but not men and victimhood. The studies so far have not been able to examine this crucial distinction. Study 6 permitted us to do so.

Study 6 also extended Study 5 by examining the possibility that perceptions of harm *intentionality* differ depending on the harmed group's gender. One aspect of moral typecasting yet to be fully examined in the current investigation is "agentic dyadic completion" (Gray & Wegner, 2010), or the tendency to perceive an intentional agent as a causal force when suffering is detected. That is, when presented with one component of the moral dyad, such as a suffering patient, individuals instinctively complete the dyad by attributing the suffering to an intentional agent (e.g., a menacing god). Dyadic

completion was tested in Study 6 by examining whether harm was assumed as more intentional when women (versus men) were in the victim role. Previous research has also demonstrated that suffering is experienced as more painful when it is intended (Gray & Wegner, 2008). Therefore, Study 6 also tested the possibility that presumed intentionality might explain the attribution of greater pain to female compared to male victims of harm. Study 6's hypotheses, methods, sample size, and analyses were pre-registered (<http://aspredicted.org/blind.php?x=b6sn2t>).

## 9.1. Method

### 9.1.1. Participants

A sample of American residents was recruited from Amazon's Mechanical Turk to complete an online survey. After removing comprehension check ( $N = 174$ ) and attention check failures ( $N = 149$ ), our final sample comprised 1599 individuals (56% female;  $M_{\text{age}} = 36.3$  years). Participants had an average of 17.3 years of work experience and the majority (70.7%) had experience in a managerial role.

### 9.1.2. Procedure

Participants were randomly assigned to evaluate one of four possible versions of a scenario in which a senior management team decided to protect profits by laying off nine employees. Across conditions, we experimentally manipulated the gender of the laid-off employees, such that they were either all men or all women, as well as the occupational field, such that employees were nurses or chemical manufacturers. According to the Bureau of Labor Statistics (2018), nursing is comprised of 84% female employees and chemical manufacturing is comprised of 74.5% male employees. Thus, gender discrimination due to adverse impact is more plausible when male employees are fired from nursing (a female-dominated field) and female employees are fired from chemical engineering (a male-dominated field). Beyond the employees' gender and occupation, the scenarios were otherwise identical. The full scenario is available in the appendix. Participants completed the following dependent measures:

9.1.2.1. *Victimhood.* The (fe)male employees who were laid off should be considered victims (1 = Strongly Disagree; 7 = Strongly Agree).

9.1.2.2. *Victim pain.* How much pain, whether it is psychological, emotional, or physical, do you think the group of (wo)men felt when they were laid off? (1 = No Pain; 7 = Extreme Pain).

9.1.2.3. *Intentionality of harm.* Jarvis' senior management team intended to harm the (fe)male employees they laid off (1 = Strongly Disagree; 7 = Strongly Agree).

9.1.2.4. *Harm inflicted* How much harm do you think the management team inflicted on the group of (wo)men they laid off? (1 = No Harm; 7 = Extreme Harm).

9.1.2.5. *Fairness.* I believe the laid off (fe)male employees have been treated fairly (1 = Strongly Disagree; 7 = Strongly Agree).

9.1.2.6. *Managers' morality.* How moral do you think the senior management team at Jarvis is? (1 = Extremely Immoral; 7 = Extremely Moral).

## 9.2. Results

A series of 2 (employee gender) by 2 (adverse impact) between-subject ANOVAs examined responses to the scenario. To examine whether responses differed by participant gender, follow-up analyses included participant gender as a third between-subject variable.

### 9.2.1. Victimhood

Contrary to predictions, employee gender did not significantly shape victimhood perceptions,  $F(1, 1595) = 1.43, p = .231$ . However, the gender by adverse impact interaction was marginally significant,  $F(1, 1595) = 3.75, p = .053$ . In the adverse impact condition where female employees were fired from a male dominated field and systematic discrimination might be inferred, judgments of the female employees' victimhood were greater ( $M = 4.19, SE = 0.09$ ), compared to when there was no adverse impact (i.e., they were fired from a female-dominated field;  $M = 3.94, SE = 0.09$ ),  $p = .049$ . When men were fired, however, judgments of victimhood did not differ regardless of whether systematic discrimination could be inferred (i.e., adverse impact because they were fired from a female-dominated field;  $M = 4.13, SE = 0.09$ ) compared to when the inference was less plausible (i.e., no adverse impact;  $M = 4.22, SE = 0.09$ ),  $p = .449$ .

A secondary model revealed the main effect of employees' gender was moderated by participant gender,  $F(1, 1591) = 10.34, p = .001$ . Male participants showed an in-group bias, such that they perceived the male employees as greater victims ( $M = 4.22, SE = 0.10$ ) than the female employees ( $M = 3.78, SE = 0.10$ ),  $p = .002$ . Female participants perceived the fired female employees as greater victims ( $M = 4.29, SE = 0.09$ ) than the male employees ( $M = 4.14, SE = 0.08$ ), but this difference was not statistically significant,  $p = .221$ .

### 9.2.2. Victim pain

Although female employees were assumed to experience greater pain ( $M = 6.03, SE = 0.03$ ) than male employees ( $M = 5.96, SE = 0.03$ ), this difference was not statistically significant,  $F(1, 1595) = 2.16, p = .142$ . A significant 2-way employee gender by adverse impact condition interaction emerged,  $F(1, 1595) = 13.36, p < .001$ . When women were fired, they were assumed to suffer more pain when gender discrimination could be inferred ( $M = 6.11, SE = 0.05$ ) compared to when gender discrimination was unlikely ( $M = 5.94, SE = 0.05$ ),  $p = .007$ . When men were fired, however, they were assumed to suffer less pain when gender discrimination was probable ( $M = 5.88, SE = 0.05$ ) compared to unlikely ( $M = 6.04, SE = 0.05$ )  $p = .013$ . Although female participants assumed the fired employees suffered greater pain across conditions ( $M = 6.06, SE = 0.03$ ) than did male participants ( $M = 5.91, SE = 0.03$ ),  $F(1, 1595) = 11.52, p = .001$ , no other participant gender effects emerged.

### 9.2.3. Intentionality of harm

A significant main effect of gender condition emerged  $F(1, 1595) = 7.04, p = .008$ , such that harm to women was assumed as more intentional ( $M = 2.52, SE = 0.05$ ) than equivalent harm to men ( $M = 2.32, SE = 0.05$ ). A significant effect of adverse impact also emerged,  $F(1, 1595) = 15.54, p < .001$ , such that harm was perceived as more intentional when discrimination seemed more plausible (i.e., high adverse impact;  $M = 2.57, SE = 0.05$ ), compared to when it was less likely (i.e., low adverse impact;  $M = 2.27, SE = 0.05$ ). However, these two main effects were qualified by a significant gender by adverse impact interaction,  $F(1, 1595) = 7.19, p = .007$ . When women were fired, their harm was perceived as more intentional when it could be due to gender discrimination ( $M = 2.77, SE = 0.08$ ), compared to when discrimination was less probable ( $M = 2.27, SE = 0.08$ ),  $p < .001$ . However, when men were fired, plausible discrimination did not significantly increase perceptions of intentional harm ( $M = 2.37, SE = 0.07$ ), compared to when discrimination was unlikely ( $M = 2.27, SE = 0.07$ ),  $p = .371$ .

The main effect of employees' gender also differed by participant gender,  $F(1, 1591) = 7.50, p = .006$ , such that female participants perceived more intentionality when women were fired ( $M = 2.61, SE = 0.07$ ) than when men were ( $M = 2.23, SE = 0.07$ ),  $p < .001$ . Male participants, on the other hand, did not perceive greater harm intentionality when the managerial team fired men ( $M = 2.44, SE = 0.08$ ) compared to when they fired women ( $M = 2.41,$

SE = 0.08),  $p = .764$ .

9.2.4. Perceived pain through harm intentionality

Although the direct effect of the employees' gender on pain was not statistically significant, previous research has demonstrated that intentional harm is perceived as more painful than unintended harm (Gray & Wegner, 2008). Because employees' gender significantly shaped perceptions of harm intentionality, we sought to examine the possibility that the female employees' pain would be perceived as (slightly) heightened because their suffering was assumed as the result of the intentional motivations of managers. Although many researchers assume that a significant direct effect (the effect from the independent variable to the dependent variable) is a necessary prerequisite of mediation, some researchers have argued that if there are theoretical reasons for expecting mediation, researchers should conduct the analysis even if there is not a significant direct effect (Rucker, Preacher, Tormala, & Petty, 2011).

We therefore constructed a mediational model using the Hayes (2018) PROCESS macro for SPSS. Indeed, a significant indirect emerged from employees' gender to assumptions of their pain through perceptions of harm intentionality,  $b = 0.012$ ,  $SE = 0.006$ ,  $CI_{95} = 0.003-0.025$  (see Fig. 3), which reduced the size of the direct effect of gender condition on pain perceptions to  $b = 0.054$ ,  $SE = 0.04$ ,  $t = 1.21$ ,  $p = .229$ .  $CI_{95} = -0.03$  to  $0.14$ . This indirect effect remained significant controlling for the adverse impact manipulation,  $b = 0.012$ ,  $SE = 0.006$ ,  $CI_{95} = 0.003-0.025$ . This partial mediation is consistent with the interpretation that victimized women are expected to experience greater pain because harm to them is perceived as more intentional than men's harm.

9.2.5. Harm inflicted

Although participants assumed the management team inflicted more harm on the fired female ( $M = 5.50$ ,  $SE = 0.05$ ) than male employees ( $M = 5.43$ ,  $SE = 0.05$ ), this difference was not statistically significant,  $F(1, 1595) = 1.16$ ,  $p = .281$ . A significant effect of the adverse impact manipulation emerged,  $F(1, 1595) = 4.67$ ,  $p = .031$ , such that participants assumed managers inflicted greater harm when discrimination was more plausible ( $M = 5.53$ ,  $SE = 0.05$ ), compared to less plausible ( $M = 5.40$ ,  $SE = 0.05$ ). These main effects were not qualified by a significant gender by discrimination condition interaction,  $F(1, 1595) = 0.41$ ,  $p = .524$ . Although female participants perceived the managers to have inflicted greater harm overall than did male participants,  $F(1, 1591) = 15.54$ ,  $p < .001$ , no other participant gender effects emerged.

9.2.6. Fairness

Perceptions that the employees were treated fairly did not differ significantly by employee gender [ $F(1, 1595) = 0.09$ ,  $p = .760$ ], plausible discrimination [ $F(1, 1595) = 2.27$ ,  $p = .132$ ], or the two-way interaction [ $F(1, 1595) = 1.81$ ,  $p = .179$ ]. Although female participants perceived the employees to be treated less fairly overall than did male participants,  $F(1, 1591) = 14.62$ ,  $p < .001$ , no other participant gender effects emerged.

9.2.7. Managers' morality

Supporting predictions, managers were perceived as less moral when they fired female employees ( $M = 3.59$ ,  $SE = 0.05$ ) compared to when they fired male employees,  $M = 3.73$ ,  $SE = 0.05$ ;  $F(1, 1595) = 4.58$ ,  $p = .033$ . A significant main effect of adverse impact also emerged,  $F(1, 1595) = 12.96$ ,  $p < .001$ , such that the managerial team was perceived as less moral when discrimination was more ( $M = 3.54$ ,  $SE = 0.05$ ) compared to less plausible ( $M = 3.78$ ,  $SE = 0.05$ ). These two main effects were not qualified by a significant gender by adverse impact interaction,  $F(1, 1595) = 0.09$ ,  $p = .765$ . Female participants perceived the managerial team as less moral overall than did male participants,  $F(1, 1591) = 5.90$ ,  $p = .015$ . No other participant gender effects emerged.

9.2.8. Perceived morality through harm intentionality

To examine the possibility that managers were evaluated as less moral when they fired women because the harm was perceived as more intentional, a mediational model was conducted using the PROCESS macro for SPSS. Indeed, a significant indirect emerged from employees' gender to perceptions of managers' morality through perceptions of harm intentionality,  $b = -0.064$ ,  $SE = 0.025$ ,  $CI_{95} = -0.11$  to  $-0.02$  (see Fig. 4), which reduced the direct effect of gender condition on managers' morality to nonsignificant,  $b = -0.076$ ,  $SE = 0.06$ ,  $t = -1.25$ ,  $p = .213$ .  $CI_{95} = -0.20$  to  $0.04$ . This indirect effect remained significant controlling for adverse impact,  $b = -0.063$ ,  $SE = 0.025$ ,  $CI_{95} = -0.11$  to  $-0.02$ . This full mediation is consistent with the interpretation that when women are in the victim role, those who inflicted this harm are perceived as agentic and intentional.

9.3. Discussion

Study 6 provided additional evidence for a gender bias in moral typecasting. As predicted by our theoretical framework, managers were perceived as more intentionally harmful and immoral when they fired a group of women than when they fired a group of men. These heightened perceptions of intentionality suggest a possible mechanism

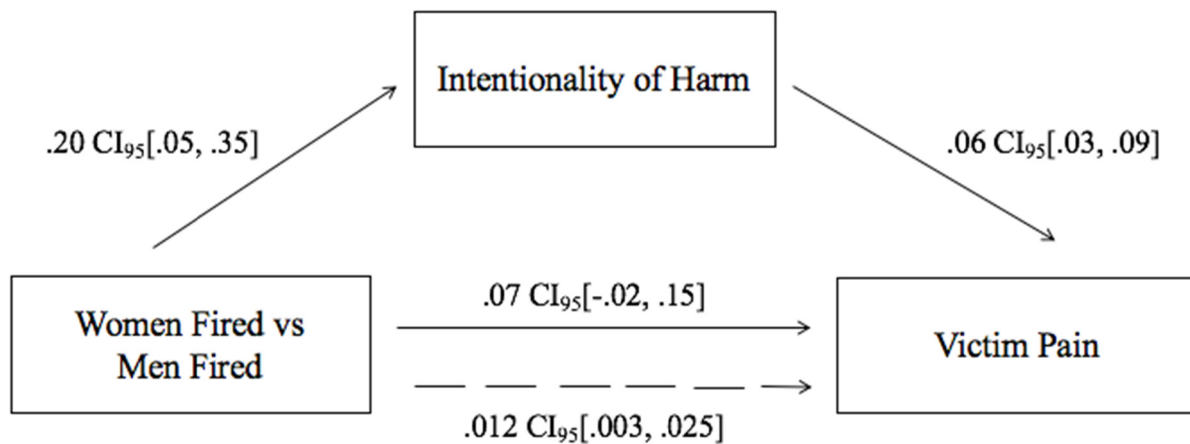
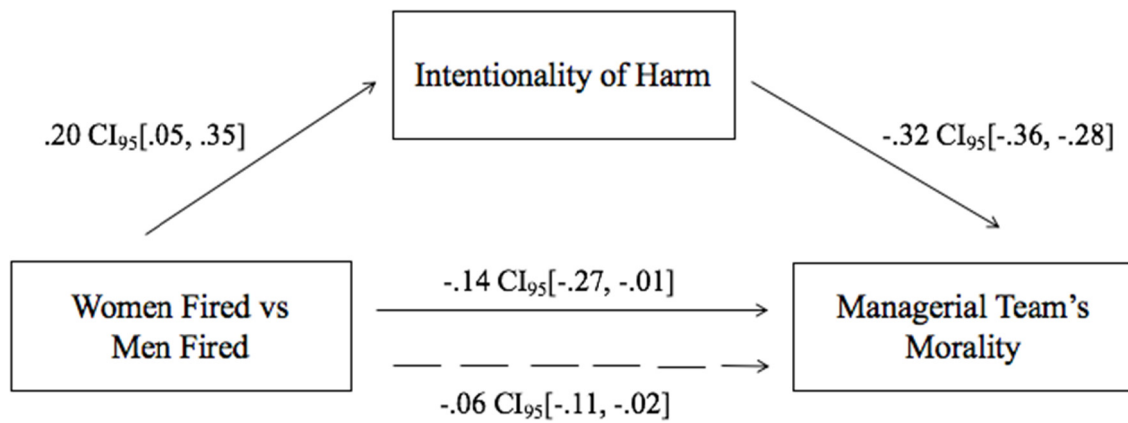


Fig. 3. Gender of Laid-off Employees Shapes Pain Perceptions Through Assumptions of Harm Intentionality (Study 6). Note: Among Study 6 participants, increased perceptions of intentionality mediated the effect of the fired employees' gender on perceptions of the victims' pain. Unstandardized coefficients depicted. Solid line from IV to DV shows the direct effect of the IV on the DV (c path); the dashed line shows the indirect effect of the IV on the DV (c-prime path).



**Fig. 4.** Gender of Laid-off Employees Shapes Evaluations of Manager's Morality Through Assumptions of Harm Intentionality (Study 6). *Note:* Among Study 6 participants, increased perceptions of intentionality mediated the effect of the fired employees' gender on evaluations of the managerial team's morality. Unstandardized coefficients depicted. Solid line from IV to DV shows the direct effect of the IV on the DV (c path); the dashed line shows the indirect effect of the IV on the DV (c-prime path).

through which those who inflict harm onto women are attributed blame and presumed to deserve punishment. Indeed, greater perceptions of harm intentionality fully mediated the relation between the fired employees' gender and assessments of the managers' morality. This pattern suggests that when victims more strongly cohere to the cognitive prototype of victim, such as when they are female, greater agency and intentionality will be attributed to perpetrators, which then influences moral judgments. This finding supports the principle of agentic dyadic completion, whereby individuals impute intentional perpetrators when they detect suffering (Gray & Wegner, 2010). Study 6 revealed nuance to dyadic completion by showing that intentionality perceptions are magnified when harmed individuals more strongly fit the cognitive prototype of victims. Likewise, perceptions of greater intentionality also mediated the effect of the employees' gender on pain perceptions. This pattern is congruent with previous findings demonstrating that intentional harm is experienced as more painful than unintentional harm (Gray & Wegner, 2008).

Results also indicate that plausible cues of gender discrimination resulting from adverse impact may amplify the typecasting of women into the victim role and the intensity of moral responses to their harm. When employees were fired from an occupation overwhelmingly comprised of the other gender, and inferences of systemic gender discrimination seemed more plausible, women, but not similarly impacted men, were more likely to be labeled victims and assumed to experience greater pain. Likewise, managers were attributed greater intentionality when women's, but not men's, firing could be attributed to gender discrimination. When male employees' firing could more (versus less) plausibly be the result of managers' discrimination, these men were assumed to experience less pain and were no more likely to be labeled victims. The findings of Study 6 suggest that when men are the possible victims of gender discrimination, this harm will be taken as incidental, rather than intentional. This interpretation is consistent with recent findings demonstrating that women's underrepresentation in various occupations is more strongly attributed to prohibitive social forces, while men's underrepresentation is attributed to internal factors, such as men's lower motivation or ability (Block, Croft, De Souza, & Schmader, 2019). The biased application of moral typecasting may contribute to this pattern, such that women are more readily identified as victims and men as agents, responsible even for their own disadvantages.

Unlike our previous studies, in-group bias manifested among both female and male participants. Consistent with the pattern of greater female in-group bias found across the other studies, female participants perceived the managers to have stronger intentions to harm the fired female (versus male) employees than male participants. However, male

participants perceived fired male employees as greater victims than fired female employees. In our six studies across 4 countries, this is the only finding showing men's reluctance to identify women as victims. Because this effect was found for only one dependent variable, it should be interpreted with caution. Men's willingness to identify male laid off employees as victims may be due to the specific context, job loss, wherein men generally experience worse outcomes compared to women (Michniewicz et al., 2014; Wang et al., 2008).

## 10. General discussion

Across six studies with a total of 3,137 participants, we found consistent support for our hypothesis that third parties exhibit a biased application of moral typecasting which cognitively links females with victimhood and males with perpetration. This pattern emerged not only with explicitly social scenarios (Studies 1, 3, 4, 5, 6), but also for animated objects devoid of humanizing attributes (Study 2). Not only did participants more easily detect female victimization and suffering, they also felt more warmly towards female versus male victims and perceived their suffering as less deserved (Study 1) and less fair (Study 5). Those who inflicted harm onto women were evaluated as more intentional and immoral (Studies 5, 6) and evoked stronger punitive responses (Study 4). These patterns suggest those who harm women will be attributed greater intentional agency and nefarious motivations because their targets more closely match the cognitive prototype of victim.

The gender bias in the application of moral typecasting also extended to the role of perpetrator. Participants more readily typecast perpetrators as male (Study 2) and desired harsher punishments for male compared to female perpetrators (Study 4). Moral outrage was amplified when a group of men was assumed to harm a group of women (Study 5), revealing the bias aggregates to group-level harm, and suggests that the combination of male perpetrators and female victims most strongly activates harm detection. Moreover, the gender bias in the evaluation of group-level harm could not be explained entirely by inferences of discrimination due to disparate impact (Study 6). Indeed, the possibility of gender discrimination enhanced the cognitive link between women and victimhood, suggesting women's harm due to systemic forces will also be more easily recognized compared to men's. The biased application of moral typecasting emerged across four cultures and persisted in circumstances where victimization was unclear, suggesting it a robust feature of human cognition. One exception to this pattern was in contexts where an individual's patency was made salient through a neurological impairment (Study 3), suggesting that third party observers do not unyieldingly apply biased typecasting, but

instead rely on more predictive cues of an individual's relative patience when evaluating harm.

We found that individuals felt less compelled to forgive and more motivated to punish male than female perpetrators, even in the form of investigations and job termination (Study 4). If this pattern extends to workplace disputes and allegations, then male employees and managers who are accused of discrimination or harassment are likely to receive disproportionately harsher penalties than female employees and managers similarly accused. Indeed, the findings of Study 4 revealed male perpetrators were less desired as coworkers and managers than female perpetrators, indicating this bias can affect employees' likelihood of ascending the organizational hierarchy, assuming they have not already lost their jobs. Research on legal sentencing converges with our findings. In the courtroom, female defendants are less likely to be found guilty and receive shorter sentences than male defendants, even accounting for crime severity (Mazzella & Feingold, 1994; Mustard, 2001; Shields & Cochran, 2019). Indeed, male offenders who victimize females receive the longest sentences, compared to other gender pairings (Curry, Lee, & Rodriguez, 2004). Taken together, these patterns suggest it is cognitively more challenging for people to detect and respond to evidence of female than male perpetration, just as it is more challenging to recognize male (versus female) victimization.

These patterns suggest that in organizations, managers and HR representatives may respond more proactively when female employees claim victimization compared to when male employees make similar claims. In the latter case, organizational decision-makers may even dispute or ignore claims of genuine male victimhood. The current findings suggest swift intervention and punishment will most likely follow female employees' claims of mistreatment from men, whereas the reverse (men claiming mistreatment by women) will be interpreted and adjudicated less seriously. If men's mistreatment is systematically met with relative apathy, this may discourage men's reporting of workplace mistreatment and discrimination, leading to underestimates of its prevalence. In recent years, organizations have implemented robust procedures to minimize gender bias in selection and promotion decisions, where greater ascriptions of male agency harm women's hierarchy ascension. As the other side of that coin, similar preemptive procedures may be necessary to minimize gender bias in responses to claims of harm, where greater ascriptions of male agency impair concern for male suffering. Organizations committed to impartiality might consider expanding bias-reducing policies from hiring and promotion contexts (e.g., blind review, independent auditing bodies) to mistreatment contexts, if the goal is to promote equal treatment of both their male and female workforce.

At first glance, this pattern of diminished punishment towards female perpetrators seems at odds with the broader literature on backlash against women's expressions of agency. However, these patterns are reconcilable. Across a range of contexts, individuals espouse the stereotype that compared to men, women are higher in communality and warmth, but lower in agency (Eagly & Mladinic, 1994; Fiske, Cuddy, & Glick, 2007). Indeed, these beliefs generally lead to more favorable views towards women (Eagly & Mladinic, 1989). Backlash occurs when an individual—of either gender—behaves unambiguously in a manner incongruent with these stereotypes. When women behave in an agentic manner, such as when they negotiate, seek power, or self-promote, they are consequently perceived as lacking in communal traits (Bowles, Babcock, & Lai, 2007; Okimoto & Brescoll, 2010; Rudman, 1998), which strongly predict interpersonal liking (Fiske, Xu, Cuddy, & Glick, 1999). Likewise, when men violate gendered stereotypes of lowered communality, such as when they advocate on behalf of others, they are consequently perceived as lacking in agency (Bosak, Kulich, Rudman, & Kinahan, 2018). These patterns suggest that agency and communality are treated, to some degree, as interrelated and oppositional traits. That is, when an individual violates their respective gender stereotype in one domain, they are assumed to also deviate from the gender stereotype in the other domain.

In the current work, we demonstrate that people continue to apply these gender stereotypes of agency and communality into contexts surrounding harm. Because men are assumed to possess greater agency, it is cognitively easier to perceive them as perpetrators than victims. Because women are assumed to possess lower agency and higher communality, it is cognitively easier to typecast them as victims than perpetrators. Moreover, the results of Study 1 suggest that when the cognitive template surrounding harm is activated (i.e., by using the language of victim and perpetrator) and the moral context is made salient, these gender stereotypes become increasingly influential. A similar interpretation can be made for Study 6, whereby increasing the plausibility of intentional harm (i.e., discrimination due to adverse impact) enhanced the cognitive link between women and victimhood. These patterns suggest that gender stereotypes might be especially likely to influence judgments in the moral domain. Perhaps when harm is at stake, human cognition reliably errors to prevent the costlier error (Haselton & Nettle, 2006). As noted in our introduction, women set the upper limit on reproduction. Thus, it is possible the human cognitive system was shaped by natural selection to make the safer error (in terms of reproduction) and protect women from harm. Likewise, if men are the primary culprits of physical violence (Archer, 2004), then it is safer to assume they are perpetrators and dole out punishments accordingly.

Our findings also appear inconsistent with the commonly held perception that female victims of sexual assault are assigned agency – and are therefore blamed – for being assaulted. Notably, research on victim blaming and punishment in the context of sexual assault has found that male sexual assault victims receive more blame and less empathy than female victims (Davies, Pollard, & Archer, 2006; Davies, Rogers, & Whitelegg, 2009; Davies, Smith, & Rogers, 2009; Gerber, Cronin, & Steigman, 2004; Osman, 2011), sexual perpetrators are more harshly blamed and punished when they victimize a woman than a man (Gerber et al., 2004; Mitchell, Angelone, Kohlberger, & Hirschman, 2009), and male sexual perpetrators receive less empathy and stronger attributions of guilt than do female perpetrators (Osman, 2011; Russell, Oswald, & Kraus, 2011). These findings parallel the pattern uncovered here and suggest that gender bias in moral typecasting extends to instances of sexual violence, but further research is required to test this possibility.

To be sure, the studies presented here have several limitations. First, they relied on self-report data and hypothetical scenarios, and cannot speak with certainty to whether actual behavior would follow the predicted pattern. However, as we report above, there is evidence for gender discrepancies in treatment across many domains that are consistent with our predictions and results. Second, our investigation only explored one boundary condition of the biased application of moral typecasting (i.e., increased patience through neurological impairment). It is possible that shifts in the other side of typecasting – increasing agency – may also minimize the gender asymmetry in typecasting. Perhaps augmenting the saliency of women's agentic capacities would eliminate the biased application of moral typecasting. If so, it is possible that female managers do not similarly benefit from the biased application of moral typecasting as women who are in non-managerial positions. Third, our investigation did not examine any of the individual differences that might predict biased application of moral typecasting beyond participant gender. Study 4's mediation analyses revealed that assumption of greater female pain predicted harsher punishments towards perpetrators. Future investigations might therefore examine whether individual differences related to pain sensitivity, such as empathy, predict the magnitude of one's gender bias in typecasting. Alternatively, endorsement of gender-related ideologies, such as feminism or benevolent sexism, may also shape the degree to which people exhibit gender asymmetries in their moral typecasting.

Four of our studies revealed female participants more readily typecast women as victims than male participants (Studies 1, 2, 5, 6). However, Study 3 found male participants were especially sensitive to female suffering. Thus, firm conclusions about which gender more



strongly exhibits biased typecasting cannot be drawn from our data. Because Study 3 employed a Chinese sample, it is possible the reversal in typecasting, such that men showed stronger asymmetries, stemmed from cultural differences. Indeed, Study 6's American male participants were more inclined to label fired male (versus female) employees victims, suggesting American men may show the predicted bias less strongly. However, given that Study 6's scenario relied on a context where men generally suffer worse outcomes—job loss (Wang et al., 2008)—it is also possible Study 6's male participants were merely more aware of these true discrepancies than were female participants.

Future research should use a broader array of scenarios and samples to determine whether men or women more strongly detect female victimization and under which circumstances. Some findings from the extant literature suggest women exhibit this biased application of typecasting more reliably than men. For example, in lab experiments, women redistributed payments more favorably toward low-earning female workers than low-earning male workers, while men showed no gender bias (Cappelen et al., 2017). This pattern was mirrored outside the laboratory; among workplace sex discrimination claims, female plaintiffs were more likely to settle and win compensation when their case was adjudicated by a female than male judge (Knepper, 2017). In online labor markets, female employers showed a stronger hiring bias in favor of female compared to matched-sample male applicants, than did male employers (Chan & Wang, 2017). In academia, female scientists believed fellow female investigators were more rational, open-minded, and demonstrated more integrity than male investigators, while male scientists did not exhibit gender-based asymmetries in their assessments (Veldkamp, Hartgerink, van Assen, & Wicherts, 2017). Our findings suggest this broader pattern may emerge from women's stronger cognitive association between women and victimhood.

Our investigation contributes to the organizational behavior literature in several ways. First, it reveals the potential for managers to systematically violate core principles of procedural justice (Leventhal, 1980; Tyler, 1994) including consistency, bias suppression, accuracy, and ethicality. Perceptions of procedural fairness have significant effects on a host of organizational outcomes (Colquitt, Conlon, Wesson, Porter, & Yee, 2001). Thus, managers who fail to apply equivalent standards of treatment to members of groups may create perceptions that organizational decisions are illegitimate, immoral, or unjust. Scholars who lament and document discrimination against women have espoused this rationale as grounds for demanding policy change to rectify such injustices. Our findings suggest these biases can cut both ways and that men's suffering may be minimized or simply ignored.

It is worth noting that our studies examined only one side of the biased application of moral typecasting. We demonstrated that in situations involving harm, women are more easily typecast as victims than men, which may lead women to receive disproportionate levels of concern or support in response to their suffering. If women are more easily categorized as patients than agents, this biased typecasting may simultaneously disadvantage women in their pursuit of leadership roles, where agency is required for reaching decisions, delegating tasks, and garnering respect (Eagly & Karau, 2002; Heilman, 2012). Indeed, a wide body of research has examined how benevolent sexism—an ideological belief that women require care and protection from men (Glick & Fiske, 1996)—harms women's access to challenging workplace opportunities that promote growth (King et al., 2012) and impairs perceptions of women's competence and suitability for managerial roles (Becker, Glick, Ilic, & Bohner, 2011). It is possible the biased application of moral typecasting demonstrated here extends to contexts that do not involve harm and contributes to reduced attributions of female agency, and consequently, leadership potential. Future research could benefit from examining the potentially far-reaching implications of this theoretical model in explaining asymmetrical treatment of men and women across various organizational contexts.

Our research also offers implications for the study of morality and behavioral ethics in organizations. As noted earlier, scholars have

documented a variety of ethical blind spots when forming moral judgments. Bazerman and Tenbrunsel (2011) argue blind spots result not only from reliance on heuristics, but also from motivated processes, such as in-group bias. These blind spots can lead individuals to behave against publicly professed values (e.g., impartiality). Our overall findings and those demonstrating women's stronger detection of female suffering (Studies 1, 2, 5, 6) suggests that in many moral dilemmas, human judgment is both a product of the application mental shortcuts (moral typecasting) and a motivated process that favors one's in-group over abstract ethical principles (e.g., every person's pain deserves equal moral consideration). Our data suggest that a complete model of how individuals fall prey to ethical blind spots should consider the characteristics of the involved actors, such as gender.

## 11. Conclusion

Although past research has examined the influence of gender stereotypes on managerial decision-making, this paper is the first, to our knowledge, to demonstrate and dissect the influence of these stereotypes in a moral context. The current findings bridge two bodies of work—descriptive gender stereotypes and moral evaluations—to reveal that gendered assumptions about relative agency or patience shape assessments of a wide and diverse set of judgments: morality, fairness, responsibility, sympathy, punishment, and compensatory aid. Although these gendered assumptions generally align with base rates of harm, possibly resembling a useful heuristic in some cases, our findings reveal that they generate systematic errors in harm evaluation. As a consequence, these biases may lead evaluators to systematically deviate from impartiality, failing to live up to the ideal that all individuals should be treated equally. The principle of impartiality is enshrined in legal systems around the world and has been propounded by civil rights activists throughout history to justify the immorality of discrimination. The biased application of moral typecasting may be used to justify unequal attributions of blame, generating disparities in punishment and aid. Crucially, the current findings suggest that *identical* transgressions will be differentially adjudicated based solely on the gender of the parties involved. Moreover, our results suggest that the moral domain is especially likely to evoke disparate gendered assumptions, indicating that some of the most egregious forms of gender discrimination may manifest in contexts involving harm.

## Funding

This work was supported by the Research Council of Norway, through its Centres of Excellence Scheme, FAIR project No 262675, and its Centres of Research-based Innovation, CSI project No 203432/030.

## CRediT authorship contribution statement

**Tania Reynolds:** Writing - original draft, Writing - review & editing, Formal analysis, Methodology. **Chuck Howard:** Writing - original draft, Methodology, Formal analysis. **Hallgeir Sjøstad:** Conceptualization, Writing-review & editing, Investigation. **Luke Zhu:** Investigation. **Tyler G. Okimoto:** Methodology, Investigation, Writing-Original draft. **Roy F. Baumeister:** Conceptualization, Funding acquisition. **Karl Aquino:** Conceptualization, Supervision, Writing - review & editing, Methodology, Funding acquisition. **JongHan Kim:** Investigation.

## Appendix

### Additional Measures and Verbatim Stimuli Across Studies

#### Study 1

A number of additional measures were captured in Study 1 for

exploratory purposes. Although beyond the scope of the current investigation, we report full details of those measures herein, as well as preliminary findings from exploratory analyses. Generally speaking, these additional results support the main finding that the assumption that the victim is female translates into more positive reactions toward the victim and more negative reactions toward the offender (irrespective of the offender's gender).

**Offender perceptions.** Participants completed the same measures in response to the offender as they did for the victim. That is, participants provided their assessments of the offender using the 10-item attitude scale ( $\alpha = 0.91$ ; Philpot & Hornsey, 2008) and the 9-item moral appraisal scale ( $\alpha = 0.83$ ; Philpot & Hornsey, 2008). They also reported their recollection of the target as being male or female (reading comprehension check). Exploratory analysis examined the association between offender gender and participants' affective responses to and moral judgments of that offender. Analysis of covariance (using label, scenario, and participant gender as covariates) indicated no significant difference of the offender gender manipulation on either affective reactions toward the offender ( $M = 2.38$ ,  $SD = 1.10$ ),  $F(1,295) = 0.04$ ,  $p = .844$ ,  $\eta^2_p < .001$ , nor moral judgements of the offender ( $M = 3.10$ ,  $SD = 1.05$ ),  $F(1,295) = 1.49$ ,  $p = .223$ ,  $\eta^2_p = .005$ .

**Willingness to integrate.** Participants were asked about their behavioral orientation toward the victim and the offender using two different scales. The first involved a 12-item behavioral orientation scale (victim  $\alpha = 0.95$ ; offender  $\alpha = 0.91$ ; Cuddy, Fiske, & Glick, 2007). The second involved a 5-item scale assessing willingness to work with the targets (victim  $\alpha = 0.96$ ; offender  $\alpha = 0.94$ ; Gromet & Okimoto, 2014). Exploratory analyses tested the effect of offender gender and assumed victim gender (and their interaction) on willingness to integrate each party (ANCOVA, with label, scenario, and participant gender as covariates). For the victim, results indicated a more positive behavioral orientation toward the victim when they assumed that victim was female ( $M_{estimated} = 6.01$ ,  $SE = 0.068$ ) versus male ( $M_{estimated} = 5.60$ ,  $SE = 0.134$ ),  $F(1,293) = 7.332$ ,  $p = .007$ ,  $\eta^2_p = .024$ . However, there was no significant effect of offender gender, or the interaction between offender and victim gender,  $F_s < 0.12$ . This pattern did not replicate when examining participants' willingness to work with the victim,  $F_s < 1.67$ . For the offender, results indicated a marginally less positive behavioral orientation toward the offender when they assumed the victim was female ( $M_{estimated} = 2.77$ ,  $SE = 0.073$ ) versus male ( $M_{estimated} = 3.09$ ,  $SE = 0.144$ ),  $F(1,293) = 3.75$ ,  $p = .054$ ,  $\eta^2_p = .013$ ; however, ratings were unaffected by offender gender or the interaction term,  $F_s < 1.5$ . Similarly, participants were less willing to work with the offender when the assumed the victim was female ( $M_{estimated} = 1.68$ ,  $SE = 0.077$ ) versus male ( $M_{estimated} = 2.12$ ,  $SE = 0.151$ ),  $F(1,293) = 6.61$ ,  $p = .011$ ,  $\eta^2_p = .022$ ; again, ratings were unaffected by offender gender or the interaction term,  $F_s < 0.65$ . Overall, participants' willingness to integrate the two parties were unaffected by offender gender, but were influenced by the assumption that the victim was female (versus male), generating more favorable reactions toward the victim and less favorable reactions toward the offender.

**Victim entitlements.** Participants were asked a series of ad hoc questions intended to tap into perceptions of the victim's entitlements ( $\alpha = 0.89$ ) by rating their 7-point scale agreement with 6 statements: 1) Organizational leaders have an obligation to help [the target]; 2) It should be made clear that there is a high degree of concern for [the target]; 3) [The target] is entitled to any type of support that he/she wants.; 4) [The target] should know that he/she was not at fault in any way; 5) Organizational leaders are obliged to acknowledge [the target]'s suffering; and 6) It should be made clear that what happened to [the target] was wrong. Exploratory analysis tested the effect of offender gender and assumed victim gender (and their interaction) on victim entitlement (ANCOVA, with label, scenario, and participant gender as covariates). However, there were no significant effects of offender gender, assumed victim gender, or their interaction when examining victim entitlement,  $F_s < 2.63$ .

**Offender punishment.** Participants were asked two questions ( $r = 0.71$ ) about offender punishment (based on Okimoto & Gromet, 2016, Study 1B): 1) How severely should [the target] be punished for his/her actions in this case? (1 = no punishment, 7 = very severe punishment); 2) What should be the maximum consequence be for [the target]'s actions? (1 = Nothing, 2 = Informal discussion, 3 = Write-up & formal discussion, 4 = Training program, 5 = Put on probation, 6 = Training + probation, 7 = Suspended without pay, 8 = Demoted, 9 = Fired from job, 10 = Fired + criminal charges). Participants were also asked a series of ad hoc questions about the offender's deservingness of punishment ( $\alpha = 0.94$ ) by rating their 7-point scale agreement with 6 statements: 1) Organizational leaders have an obligation to punish [the target]; 2) [The target] should be made to suffer for his/her actions; 3) [The target] deserves to be punished; 4) [The target] needs to be taught a lesson; 5) Organizational leaders should punish [the target]; and 6) [The target] ought to suffer for what s/he did to [the victim]. Exploratory analysis tested the effect of offender gender and assumed victim gender (and their interaction) on all three measures capturing offender punishment (ANCOVA, with label, scenario, and participant gender as covariates). Results indicated no significant effects,  $F_s < 2.28$ .

**Empathy.** Participants reported their level of cognitive and emotional empathy (Davis, 1983) felt toward both the victim (cognition  $\alpha = 0.85$ ; emotion  $\alpha = 0.90$ ) and offender (cognition  $\alpha = 0.91$ ; emotion  $\alpha = 0.79$ ). Exploratory analysis tested the effect of offender gender and assumed victim gender (and their interaction) on victim/offender cognitive/emotional empathy (ANCOVA, with label, scenario, and participant gender as covariates). For empathy toward the victim, results indicated no significant effects,  $F_s < 1.89$ . For empathy toward the offender, results showed less emotional empathy toward the offender when the victim was assumed to be female ( $M_{estimated} = 1.79$ ,  $SE = 0.065$ ) versus male ( $M_{estimated} = 2.22$ ,  $SE = 0.128$ ),  $F(1,293) = 7.226$ ,  $p = .008$ ,  $\eta^2_p = .024$ . However, there was no parallel effect on cognitive empathy,  $F(1,293) = 0.129$ ,  $p = .720$ ,  $\eta^2_p < .001$ . As before, there were no significant effects of offender gender or the interaction between offender and victim gender on empathy toward the offender,  $F < 1.11$ .

**Restorative/retributive orientation.** At the end of the survey, participants self-reported their justice orientation with a revised version of the measure from Okimoto, Wenzel, and Feather (2012). This measure was included for scale validation purposes, and was not relevant to the current investigation; therefore, no exploratory analyses were conducted.

### Study 3

The scenario used in Study 3 is presented below. Depending on condition, Employee A and B were named either Carrie or Steve. *Italicized* statements were presented for the neurological impairment condition only.

*"[Employee A] has an untreatable neurological condition and a slight cognitive impairment.* [Employee A] is very sociable and often approaches people (s)he doesn't know and starts talking to them. *However, (s)he has difficulty reading social cues and often doesn't know how to act appropriately, especially in professional environments.*

[Employee A] works in an office with [Employee B]. [Employee A] is fairly new to the job and wants to make friends, so (s)he goes out of his/her way to interact with the people (s)he meets. For example, (s)he often high-fives or fist bumps co-workers as (s)he passes them in the halls. Sometimes, (s)he hugs people (s)he particularly likes or touches their arm when talking to them. [Employee A] also likes to pay compliments, but his/her compliments often seem odd to their recipients. For example, (s)he told one co-worker, "You're a lot smarter than you seem."

[Employee A] likes [Employee B]. [Employee A]'s not sexually attracted to him/her, but (s)he thinks (s)he is nice and often stops by

his/her cubicle during the day to say hello. [Employee B] finds [Employee A]’s interest in him/her distracting. [Employee B] once had a bad experience with [Employee A] when (s)he gave him/her a high five in the lunchroom. [Employee A] had mayonnaise on his/her hand from a sandwich (s)he was eating, which got onto [Employee B]’s hand, and slightly repulsed him/her. After that, (s)he started dreading having to interact with [Employee A].

One day [Employee B] was late for work because his/her car broke down and it was raining. [Employee B] had to pay \$90 to have his/her car towed to the shop and (s)he was soaking wet when (s)he got to work. On his/her way to his/her cubicle, (s)he happened to run into [Employee A]. [Employee A] asked him/her what happened and why (s)he was so wet. [Employee B] told him/her about the car. [Employee A] then drew him/her towards him/her for a hug and said “Poor [Employee B]. I hope the rest of your day goes better.” [Employee B] tensed up and after (s)he let him/her go, (s)he told him/her (s)he was really busy and needed to get back to work. [Employee B] didn’t see him/her again for the rest of the day.

Throughout the week, [Employee B] thought about [Employee A]’s behaviors and decided (s)he didn’t want him/her around because (s)he made him/her uncomfortable. [Employee B] wrote a report listing all the ways (s)he had created an uncomfortable work environment for him/her, including touching him/her inappropriately. [Employee B] sent the report to the company’s human resources manager. A couple of days later, [Employee A] was fired.”

#### Dependent measures

**Perceived control.** Participants indicated the degree to which they perceived the fired employee to have control over his/her actions (1 = *no control at all*; 7 = *absolute control*). An independent samples *t*-test revealed perceptions of control did not differ based on the neurological impairment manipulation,  $t(210) = -1.16, p = .245$ .

#### Study 4

**Harassment.** Five items assessed whether participants perceived the comment as harassment: 1) [Commenter] told a sexually explicit comment; 2) [Commenter]’s sexual comment should be considered sexual harassment; 3) Do you think [Recipient] would consider [Commenter]’s comment funny? (reverse coded); 4) Do you believe [Recipient] would feel insulted by [Commenter]’s comment?; 5) Do you believe [Commenter] intended to sexually harass [Recipient]? ( $\alpha = 0.84$ ).

Main effects of recipient gender [ $F(1, 210) = 10.79, p = .001$ ] and commenter gender [ $F(1, 210) = 5.63, p = .019$ ] were qualified by an interaction between these two variables [ $F(1, 210) = 4.63, p = .033$ ]. Consistent with the proposition that men are more easily typecast as agents and women as victims, contrast analysis revealed that participants perceived the joke to be more harassing when the recipient was female and the commenter was male ( $M = 3.53, SE = 0.11$ ) than when the recipient was male and the commenter was female ( $M = 2.90, SE = 0.12, t(210) = 4.00, p < .001$ ). Perceptions of harassment did not differ between the latter condition and the two same-sex conditions ( $p$ ’s  $> 0.42$ ).

**Reporting the comment.** To measure the degree to which participants felt the comment deserved to be reported they completed two items: 1) If you were [Recipient], how likely would you be to report Commenter’s behavior to his/her supervisor?; 2) I think [Recipient] would be totally justified if (s)he went to Commenter’s supervisor and accused Commenter of sexual harassment. ( $\alpha = 0.83$ ).

Consistent with the proposition that women more easily typecast into the role of a victim who requires protection, a main effect of recipient gender,  $F(1, 210) = 11.39, p = .001$ , revealed participants believed the joke should be reported when it was directed at a woman ( $M = 2.94, SE = 0.11$ ) more than when directed at a man ( $M = 2.44, SE = 0.10$ ). Furthermore, a main effect of commenter gender,  $F(1,$

$210) = 5.39, p = .021$ , revealed participants believed the joke more strongly warranted reporting when it was delivered by a man ( $M = 2.87, SE = 0.10$ ) than by a woman ( $M = 2.52, SE = 0.11$ ). This is consistent with the proposition that men are more easily typecast into the role of a perpetrator deserving punishment. There was no significant interaction between commenter and recipient gender,  $F(1, 210) = 0.67, p = .41$ .

**Sexual intentions.** To measure the extent to which participants perceived the commenter to be sexually aroused we asked ‘Do you think that A feels sexually aroused by seeing B bend over to pick up the fork?’ and ‘Do you think that seeing B in the bent over position makes him/her an object of sexual desire for A?’ ( $\alpha = 0.88$ ). Main effects of recipient gender,  $F(1, 210) = 27.17, p < .001$ , and commenter gender,  $F(1, 210) = 13.89, p < .001$  were qualified by an interaction between these two variables,  $F(1, 210) = 13.63, p < .001$ . Contrast analysis revealed participants believed the commenter was more sexually aroused in the male commenter-female recipient condition ( $M = 3.31, SE = 0.14$ ) than when both parties were female [ $M = 2.41, SE = 0.12; t(210) = 5.25, p < .001$ ], both parties were male [ $M = 2.23, SE = 0.12, t(210) = 6.27, p < .001$ ], or when the commenter was female and the recipient was male [ $M = 2.22, SE = 0.11, t(210) = 6.32, p < .001$ ]. Perceptions of sexual arousal did not differ between the latter three dyads ( $p$ ’s  $> 0.28$ ).

**Character.** We measured participants’ judgments of the commenter’s character along four dimensions (Goodwin, Piazza, & Rozin, 2014). To measure high morality-high warmth (HMHW), we used the items: unkind (reverse coded), not helpful (reverse coded), generous, caring, compassionate, and cooperative ( $\alpha = 0.90$ ). To measure high morality-low warmth (HMLW) we used the items: untrustworthy (reverse coded), sincere, responsible, and fair ( $\alpha = 0.86$ ). To measure low morality-high warmth (LMHW), we used the items: hostile (reverse coded), warm, outgoing, and friendly ( $\alpha = 0.77$ ). To measure ability we used the items: hardworking, competent, and intelligent ( $\alpha = 0.89$ ).

Perceptions of the commenter’s character along the high morality-high warmth (HMHW) and high morality-low warmth (HMLW) dimensions followed the same pattern. There was a main effect of commenter gender,  $F(1, 210) = 5.42, p = .021$  for HMHW;  $F(1, 210) = 4.60, p = .033$  for HMLW, such that male commenters were rated lower on these dimensions than female commenters (mean difference for HMHW =  $-0.24, p = .021$ ; mean difference for HMLW =  $-0.23, p = .034$ ). There was no main effect of recipient gender,  $F(1, 210) = 0.34, p = .56$  for HMHW;  $F(1, 210) = 0.87, p = .35$  for HMLW, and there was no commenter by recipient gender interaction,  $F(1, 210) = 2.59, p = .11$  for HMHW;  $F(1, 210) = 1.99, p = .16$  for HMLW.

Turning to perceptions of the commenter’s character along the LWHM dimension, there was a marginal main effect of commenter gender,  $F(1, 210) = 3.67, p = .057$ , no main effect of recipient gender,  $F(1, 210) = 0.73, p = .39$ , and a significant commenter by recipient gender interaction,  $F(1, 210) = 4.87, p = .028$ . Contrast analysis revealed participants rated the female commenter who targeted a man higher on the LMHW dimension ( $M = 3.62, SE = 0.09$ ) than the female commenter who targeted another female [ $M = 3.31, SE = 0.11, t(210) = 2.18, p = .031$ ], the male commenter who targeted another male [ $M = 3.20, SE = 0.12, t(210) = 2.91, p = .004$ ], and the male commenter who target a female [ $M = 3.34, SE = 0.09, t(210) = 1.96, p = .051$ ]. Perceptions did not differ between the latter three dyads ( $p$ ’s  $> 0.34$ ).

Last, with respect to perceptions of the commenter’s ability, there was a main effect of commenter gender,  $F(1, 210) = 5.74, p = .017$ , a marginal main effect of recipient gender,  $F(1, 210) = 3.35, p = .069$ , and a marginal commenter by recipient gender interaction,  $F(1, 210) = 3.35, p = .069$ . Contrast analysis revealed participants rated the female commenter joking with a man higher in ability ( $M = 3.64, SE = 0.10$ ) than the female commenter joking with another female [ $M = 3.22, SE = 0.11, t(210) = 2.60, p = .010$ ], the male commenter

joking with another male [ $M = 3.15$ ,  $SE = 0.13$ ,  $t(210) = 2.99$ ,  $p = .003$ ], and the male commenter joking with a female [ $M = 3.15$ ,  $SE = 0.09$ ,  $t(210) = 2.99$ ,  $p = .003$ ]. Perceptions did not differ between the latter three dyads ( $p$ 's  $> 0.69$ ).

**Sexual objectification beliefs.** To measure differences in beliefs about sexual objectification we asked 'Is it wrong for a man/woman to treat a man/woman as an object of sexual desire?' and 'Do you think that a man/woman treating a man/woman as a sexual object is an act of oppression?' where man/woman matched the gender of the commenter and recipient in the condition ( $\alpha = 0.84$ ). There was no main effect of commenter gender,  $F(1, 210) = 2.52$ ,  $p = .11$ , and no main effect of recipient gender,  $F(1, 210) = 0.19$ ,  $p = .66$ . There was, however, a significant commenter by recipient gender interaction,  $F(1, 210) = 4.42$ ,  $p = .037$ . Contrast analysis revealed participants felt sexual objectification was significantly more oppressive and objectionable when a man objectifies a woman ( $M = 3.77$ ,  $SE = 0.18$ ) than when a woman objectifies a woman [ $M = 3.16$ ,  $SE = 0.17$ ,  $t(210) = 2.61$ ,  $p = .010$ ], marginally more oppressive and objectionable when a man objectifies a man [ $M = 3.35$ ,  $SE = 0.16$ ,  $t(210) = 1.79$ ,  $p = .075$ ], and directionally more oppressive and objectionable when a woman objectifies a man [ $M = 3.44$ ,  $SE = 0.16$ ,  $t(210) = 1.43$ ,  $p = .154$ ]. Beliefs regarding objectification did not differ between the latter three dyads ( $p$ 's  $> 0.23$ ).

**Attractiveness.** Participants provided their perceptions of the targets' attractiveness by answering the following questions on a 7-point Likert scale (1 = *Very unattractive*, 7 = *Very attractive*): 'How would you rate A in terms of physical attractiveness?' and 'How would you rate B in terms of physical attractiveness?' These items were included to explore the idea that perceived attractiveness of a commenter or recipient might vary based on the gender of the other person in the dyad. More concretely, this measure allowed us to test the hypotheses that a male commenter is seen as less attractive when he tells the joke to a woman compared to when he tells it to a man, and that a female recipient is seen as less attractive when targeted by a man vs. woman. These hypotheses were not supported ( $p$ 's  $> 0.24$ ).

## Study 5

Below we present the scenario used in Study 5:

"The following is a case involving a company and a decision it made about how to maintain its profitability. Please read the case carefully as you will be asked answer several questions about the company's actions.

### Jarvis manufacturing

Jarvis Manufacturing is a private company that makes specialized ski apparel. It employs 200 people and has been in business for 30 years. The company's profits have been slowly declining as the market has become saturated with competitors from Eastern Europe and Asia. Although the company is not in serious financial trouble yet, its top management team, which consists of the company's founder and CEO, Chief Operating and Finance Officer, and Vice President of Human Resources, analyzed the company's structure, operations, and financial situation at a recent strategy meeting. They concluded that one way Jarvis can maintain its profitability and remain competitive is to lay off some of its workforce.

The jobs they decided to eliminate were redundant and could be performed by fewer people without seriously compromising company operations. Furthermore, the management team estimated that eliminating these jobs would increase profitability by about 3–5% per year over the next three years. However, some of the employees currently in these jobs had been working at Jarvis for some time. In its final analysis, the management team believed that eliminating these jobs was necessary to streamline operations and improve the company's position in an increasingly competitive marketplace.

In the end, the management team implemented their plan and 9 employees were laid off. **All of them were men/women** and none of them had college degrees. They had been working at Jarvis for an average of 5 years.

### The aftermath

Jarvis gave each man/women they laid off a severance package of two week's salary and \$100 for each year they had been employed with the company. Three months later, 7 of the 9 men/women who had been laid off found new jobs, but at lower salaries than they received at Jarvis. Three of the men/women had to relocate to a new city to get a job. Two were still unemployed.

One year after the layoffs, Jarvis increased its profit margin by 3% and operations were running more efficiently."

## Study 6

Below is the scenario used in Study 6, with portions varying across condition underlined:

"Jarvis is a private company that specializes in chemical manufacturing/nursing care. It employs 200 manufacturing/nursing staff and has been in business for 30 years. The company's profits have been slowly declining as the market has become saturated with competitors. In a recent strategy meeting, the top management team (the CEO, Chief Operating Officer, Chief Financial Officer) analyzed Jarvis' structure, operations, and financial situation. They concluded that one way Jarvis can increase its profitability and remain competitive is to lay off some of its workforce.

The management team decided to eliminate jobs that were redundant and could be performed by fewer people without seriously compromising company operations. Furthermore, the management team estimated that eliminating these jobs would increase profitability by about 3–5% per year over the next three years. However, some of the employees currently in these jobs had been working at Jarvis for some time. In its final analysis, the management team believed that eliminating these jobs was necessary to streamline operations and improve the company's position in an increasingly competitive marketplace.

In the end, the management team implemented their plan and 9 manufacturing/nursing employees were laid off. All of them were men/women, and each had been with the company for around 5 years. None of these men/women had college degrees, which would make securing a new position in an increasingly competitive workforce challenging."

## References

- Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology*, 85, 768–776.
- Archer, J. (2004). Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology*, 8, 291–322.
- Bauman, A. E., Reis, R. S., Sallis, J. F., Wells, J. C., Loos, R. J., Martin, B. W., & Lancet Physical Activity Series Working Group. (2012). Correlates of physical activity: Why are some people physically active and others not?. *The Lancet*, 380, 258–271.
- Bazerman, M. H., & Tenbrunsel, A. E. (2011). *Blind spots: Why we fail to do what's right and what to do about it*. Princeton, NJ: Princeton University Press.
- Becker, J. C., Glick, P., Ilic, M., & Bohner, G. (2011). Damned if she does, damned if she doesn't: Consequences of accepting versus confronting patronizing help for the female target and male actor. *European Journal of Social Psychology*, 41, 761–773.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, 155–162.
- Berry, D. S., & McArthur, L. Z. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology*, 48, 312–323.
- Block, K., Croft, A., De Souza, L., & Schmader, T. (2019). Do people care if men don't care about caring? The asymmetry in support for changing gender roles. *Journal of Experimental Social Psychology*, 83, 112–131.
- Bosak, J., Kulich, C., Rudman, L., & Kinahan, M. (2018). Be an advocate for others, unless you are a man: Backlash against gender-atypical male job candidates. *Psychology of Men & Masculinity*, 19, 156–165.
- Bosson, J. K., Vandello, J. A., Burnaford, R. M., Weaver, J. R., & Arzu Wasti, S. (2009). Precarious manhood and displays of physical aggression. *Personality and Social Psychology Bulletin*, 35, 623–634.

- Bowlby, J. (1969). *Attachment and Loss*, Vol. I. New York: Basic Books.
- Bowles, H. R., Babcock, L., & Lai, L. (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes*, 103, 84–103.
- Burnstein, E., Crandall, C., & Kitayama, S. (1994). Some neo-Darwinian decision rules for altruism: Weighing cues for inclusive fitness as a function of the biological importance of the decision. *Journal of Personality and Social Psychology*, 67, 773–789.
- Cappelen, A., Falch, R., & Tungodden, B. (2017). *The boy crisis: Experimental evidence on the acceptance of males falling behind. Discussion paper 21/2017*. Norway: Norwegian School of Economics.
- Chan, J., & Wang, J. (2017). Hiring preferences in online labor markets: Evidence of a female hiring bias. *Management Science*.
- Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86, 425–445.
- Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. *Journal of Personality and Social Psychology*, 50, 925–935.
- Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. H. (1995). “Their ideas of beauty are, on the whole, the same as ours”: Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology*, 68, 261–279.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 92, 631–648.
- Curry, T. R., Lee, G., & Rodriguez, S. F. (2004). Does victim gender increase sentence severity? Further explorations of gender dynamics and sentencing outcomes. *Crime & Delinquency*, 50, 319–343.
- Davies, M., Pollard, P., & Archer, J. (2006). Effects of perpetrator gender and victim sexuality on blame toward male victims of sexual abuse. *The Journal of Social Psychology*, 146, 275–291.
- Davies, M., Rogers, P., & Whitelegg, L. (2009). Effects of victim gender, victim sexual orientation, victim response and respondent gender on judgements of blame in a hypothetical adolescent rape. *Legal and Criminological Psychology*, 14, 331–338.
- Davies, M., Smith, R., & Rogers, P. (2009). Police perceptions of rape as function of victim gender and sexuality. *The Police Journal*, 82, 4–12.
- Dijker, A. J. (2001). The influence of perceived suffering and vulnerability on the experience of pity. *European Journal of Social Psychology*, 31, 659–676.
- Dijker, A. J. (2010). Perceived vulnerability as a common basis of moral emotions. *British Journal of Social Psychology*, 49, 415–423.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109, 573–598.
- 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., & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *European Review of Social Psychology*, 5, 1–35.
- FeldmanHall, O., Dalgleish, T., Evans, D., Navrady, L., Tedeschi, E., & Mobbs, D. (2016). Moral chivalry: Gender and harm sensitivity predict costly altruism. *Social Psychological and Personality Science*, 7, 542–551.
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11, 77–83.
- Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis) respecting versus (dis) liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues*, 55, 473–489.
- Gerber, G. L., Cronin, J. M., & Steigman, H. J. (2004). Attributions of blame in sexual assault to perpetrators and victims of both genders. *Journal of Applied Social Psychology*, 34, 2149–2165.
- Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70, 491–512.
- Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: An evolutionary analysis and empirical review. *Psychological Bulletin*, 136, 351–374.
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106, 148–168.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315, 619.
- Gray, K., Waytz, A., & Young, L. (2012). The moral dyad: A fundamental template unifying moral judgment. *Psychological Inquiry*, 23, 206–215.
- Gray, K., & Wegner, D. M. (2008). The sting of intentional pain. *Psychological Science*, 19, 1260–1262.
- Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, 96, 505–520.
- Gray, K., & Wegner, D. M. (2010). Blaming God for our pain: Human suffering and the divine mind. *Personality and Social Psychology Review*, 14, 7–16.
- Gray, K., & Wegner, D. M. (2011). To escape blame, don't be a hero—Be a victim. *Journal of Experimental Social Psychology*, 47, 516–519.
- Haidt, J. (2003). The moral emotions. *Handbook of Affective Sciences*, 11, 852–870.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Pantheon.
- Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review*, 10, 47–66.
- Hauser, M. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York: HarperCollins Publishers.
- Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition: A Regression-Based Approach*. New York, New York: The Guilford Press.
- Hayes, A. F. (2018). Partial, conditional, and moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, 85, 4–40.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57, 243–259.
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57, 657–674.
- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, 113–135.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- Keating, C. F., Randall, D. W., Kendrick, T., & Gutshall, K. A. (2003). Do babyfaced adults receive more help? The (cross-cultural) case of the lost resume. *Journal of Nonverbal Behavior*, 27, 89–109.
- King, E. B., Botsford, W., Hebl, M. R., Kazama, S., Dawson, J. F., & Perkins, A. (2012). Benevolent sexism at work: Gender differences in the distribution of challenging developmental experiences. *Journal of Management*, 38, 1835–1866.
- Knepper, M. (2017). When the shadow is the substance: Judge gender and the outcomes of workplace sex discrimination cases. *Journal of Labor Economics*.
- Kobrynowicz, D., & Branscombe, N. R. (1997). Who considers themselves victims of discrimination? *Psychology of Women Quarterly*, 21, 347–363.
- Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, 137, 616–642.
- Krebs, D. L. (2008). Morality: An evolutionary account. *Perspectives on Psychological Science*, 3, 149–172.
- Lassek, W. D., & Gaulin, S. J. (2009). Costs and benefits of fat-free muscle mass in men: Relationship to mating success, dietary requirements, and native immunity. *Evolution and Human Behavior*, 30, 322–328.
- Leventhal, G. S. (1980). What should be done with equity theory? New approaches to the study of fairness in social relationships. In K. Gergen, M. Greenberg, & R. Willis (Eds.), *Social exchange: Advances in theory and research* (pp. 27–55). New York: Plenum.
- Mazzella, R., & Feingold, A. (1994). The effects of physical attractiveness, race, socio-economic status, and gender of defendants and victims on judgments of mock jurors: A meta-analysis. *Journal of Applied Social Psychology*, 24, 1315–1338.
- Michniewicz, K. S., Vandello, J. A., & Bosson, J. K. (2014). Men's (mis) perceptions of the gender threatening consequences of unemployment. *Sex Roles*, 70, 88–97.
- Mitchell, D., Angelone, D. J., Kohlberger, B., & Hirschman, R. (2009). Effects of offender motivation, victim gender, and participant gender on perceptions of rape victims and offenders. *Journal of Interpersonal Violence*, 24, 1564–1578.
- Mustard, D. B. (2001). Racial, ethnic, and gender disparities in sentencing: Evidence from the US federal courts. *The Journal of Law and Economics*, 44, 285–314.
- Nosek, B., & Banaji, M. B. (2002). The go/no-go association task. *Social Cognition*, 19, 625–664.
- Okimoto, T. G., & Brescoll, V. L. (2010). The price of power: Power seeking and backlash against female politicians. *Personality and Social Psychology Bulletin*, 36, 923–936.
- Osman, S. L. (2011). Predicting rape empathy based on victim, perpetrator, and participant gender, and history of sexual aggression. *Sex Roles*, 64, 506–515.
- Philpot, C. R., & Hornsey, M. J. (2008). What happens when groups say sorry: The effect of intergroup apologies on their recipients. *Personality and Social Psychology Bulletin*, 34, 474–487.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879–891. <https://doi.org/10.3758/BRM.40.3.879>.
- Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2013). *HLM 7.01 for Windows [Computer software]*. Skokie, IL: Scientific Software International Inc.
- Richeson, J. A., & Ambady, N. (2001). Who's in charge? Effects of situational roles on automatic gender bias. *Sex Roles*, 44, 493–512.
- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 5, 359–371.
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, 74, 629–645.
- Rudman, L. A., & Goodwin, S. A. (2004). Gender differences in automatic in-group bias: Why do women like women more than men like men? *Journal of Personality and Social Psychology*, 87, 494–509. <https://doi.org/10.1037/0022-3514.87.4.494>.
- 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.
- Rudolph, U., Roesch, S., Greitemeyer, T., & Weiner, B. (2004). A meta-analytic review of help giving and aggression from an attributional perspective: Contributions to a general theory of motivation. *Cognition and Emotion*, 18, 815–848.
- Russell, B. L., Oswald, D. L., & Kraus, S. W. (2011). Evaluations of sexual assault: Perceptions of guilt and legal elements for male and female aggressors using various coercive strategies. *Violence and Victims*, 26, 799–815.
- Schmitt, M. T., Branscombe, N. R., Kobrynowicz, D., & Owen, S. (2002). Perceiving discrimination against one's gender group has different implications for well-being in women and men. *Personality and Social Psychology Bulletin*, 28, 197–210.
- Schwartz, S. H., & Rubel, T. (2005). Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of Personality and Social Psychology*, 89, 1010–1028.
- Shields, R. T., & Cochran, J. C. (2019). The gender gap in sex offender punishment. *Journal of Quantitative Criminology*, 1–24.
- Sjåstad, H., & Baumeister, R. F. (2018). The future and the will: Planning requires self-control, and ego depletion leads to planning aversion. *Journal of Experimental Social Psychology*, 76, 127–141.
- Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal*

- of *Personality and Social Psychology*, 48, 813–838.
- Tanner, J. M. (1978). *Foetus into man: Physical growth from conception to maturity*. London: Open Books.
- Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics*. New York, NY: W. W. Norton & Company.
- Tyler, T. R. (1994). Psychological models of the justice motive. *Journal of Personality and Social Psychology*, 67, 850–863.
- Veldkamp, C. L., Hartgerink, C. H., van Assen, M. A., & Wicherts, J. M. (2017). Who believes in the storybook image of the scientist? *Accountability in Research*, 24, 127–151.
- Wang, J., Lesage, A., Schmitz, N., & Drapeau, A. (2008). The relationship between work stress and mental disorders in men and women: Findings from a population-based study. *Journal of Epidemiology & Community Health*, 62, 42–47.
- Weiner, B. (1980). A cognitive (attribution)-emotion-action model of motivated behavior: An analysis of judgments of help-giving. *Journal of Personality and Social Psychology*, 39, 186–200.
- Workplace Bullying Institute (2017). Women still majority targets of abusive conduct at

work: 2017 WBI US survey. Retrieved from <https://www.workplacebullying.org/gender/>.

### Appendix References

- Davis, M. (1983). Measuring individual differences in empathy: Evidence for a multi-dimensional approach. *Journal of Personality and Social Psychology*, 44, 113–126.
- Gromet, D. M., & Okimoto, T. G. (2014). Back into the fold: The influence of offender amends and victim forgiveness on reintegration. *Business Ethics Quarterly*, 24, 411–441.
- Okimoto, T. G., & Gromet, D. M. (2016). Differences in sensitivity to deviance partly explain ideological divides in social policy support. *Journal of Personality and Social Psychology*, 111, 98–117.
- Okimoto, T. G., Wenzel, M., & Feather, N. T. (2012). Retribution and restoration as general orientations toward justice. *European Journal of Personality*, 26, 255–275.