

DOES GENDER RAISE THE ETHICAL BAR?
EXPLORING THE PUNISHMENT OF ETHICAL VIOLATIONS AT WORK

Jessica A. Kennedy
Owen Graduate School of Management
Vanderbilt University
jessica.kennedy@owen.vanderbilt.edu

Mary-Hunter McDonnell
The Wharton School
University of Pennsylvania
marymcd@wharton.upenn.edu

Nicole M. Stephens
Kellogg School of Management
Northwestern University
n-stephens@kellogg.northwestern.edu

Please direct correspondence to:

Jessica A. Kennedy
401 21st Ave. S
Nashville, TN 37203
jessica.kennedy@owen.vanderbilt.edu

ABSTRACT

We investigate whether women are targets of more severe punishment than men following ethical violations at work. Using an experimental design, Study 1 finds evidence that ethical behavior is more strongly prescribed for women than for men, even when they occupy an identical professional role. Study 2 manipulates the gender of a manager in a hypothetical scenario and finds that women are punished more severely than men for ethical violations at work. It also tests the scope of our theory by asking whether women are punished more for errors in general, or only for intentional ethical violations. Using field data, Study 3 examines how severely attorneys are punished for violating the American Bar Association's ethical rules. Female attorneys are punished more severely than male attorneys, after accounting for a variety of factors. Greater representation of women among decision-makers diminishes the gender disparity in punishment. Our research documents a new prescriptive stereotype faced by women and helps to explain the persistence of gender disparities in organizations. It highlights punishment severity as a novel mechanism by which institutions may derail women's careers more than men's.

Keywords: Gender; Ethical Violations; Punishment; Stereotypes

Over several decades, social scientists have unveiled evidence of persistent, multi-faceted discrimination against professional women seeking to rise in the ranks of their organizations. Women face greater challenges than men at many stages of their ascent, including when they obtain higher education (Milkman, Akinola, & Chugh, 2012), seek to enter organizations (Belliveau, 2005; Goldin & Rouse, 2000; Gorman 2005), and when their performance is evaluated to determine promotion or compensation (Beckman & Phillips, 2005; Belliveau, 2012; Castilla, 2008; Castilla, 2011; DiPrete and Soule, 1988; Fernandez-Mateo, 2009; Joshi, 2014; McGinn & Milkman, 2013; Olson & Becker 1983). At each stage, gender stereotypes bias judgments and decisions, hindering women's advancement (Amanatullah & Morris, 2010; Brescoll, 2011; Heilman, 2012; Rudman & Phelan, 2008). All of these decision contexts provide evidence that gender bias shapes the allocation of employment-related *rewards*. The net effect is that women receive less compensation (Dreher & Cox, 2000) and claim less authority (Bowles & McGinn, 2005; Wright, Baxter, & Birkelund, 1995) than their performance justifies (Heilman, 2012). As a consequence, women ascend the corporate hierarchy more slowly than their male counterparts.

In this research, we investigate another form of discrimination against women professionals: how gender bias shapes the allocation of employment-related *punishment*. Specifically, we ask whether women suffer more severe punishment than men for ethical violations at work. Harsher punishment could be one important but subtle way of subduing women who have achieved inroads into prestigious professions. Lower social tolerance for gender-based discrimination and the threat of legal liability have driven blatant discrimination underground in many workplaces, but discrimination persists in subtle forms (Berdahl, 2007; Brescoll, 2011; Cortina, 2008; Dipboye & Halverson, 2004; Ely & Meyerson, 2000; Ryan & Haslam, 2007). Harsher punishment for ethical violations may be one important example of covert discrimination. As discussed, earlier research has established that it is more difficult for women to ascend in organizations. If women are punished more severely than men, they may also descend more quickly. In other words, harsher punishment could be another example of how women's professional status is more precarious than men's (Brescoll, Dawson, & Uhlmann, 2010; Ryan & Haslam, 2007).

Across three studies, we investigate the relationship between gender, the prescription to be ethical, and punishment for ethical violations. Consistent with prior research (Bowles & Gelfand, 2010; Brescoll et al., 2010; Castilla & Benard, 2010; Uhlmann & Cohen, 2005; Eagly, Makhijani, & Klonsky, 1992; Rudman, 1998; Williams & Tiedens, 2016), we expect women to be held to different standards than men. We extend prior work by examining *ethical* standards. We expect that women face an intensified prescriptive stereotype to be ethical and will be punished more severely than men for ethical violations at work. Punishment has significant implications for professionals' career outcomes (Butterfield, Trevino, Wade & Ball, 2005; Trevino, 1992). It reduces professionals' status in their field and has long-lasting effects, both psychologically (Arvey & Ivancevich, 1980) and professionally. When punished unfairly, people's job performance, commitment, and organizational citizenship behavior decline, while counterproductive behaviors increase (Cohen-Charash & Spector, 2001). If women systematically experience harsher punishment than men following ethical violations, this could help to explain well-documented disparities in pay, advancement, and job satisfaction (Bowles & McGinn, 2005; Ely, Stone, & Ammerman, 2014).

To examine the relationship between gender and punishment severity, we take a multi-method approach. We begin with an experimental survey exploring gender and the prescriptive stereotype to be ethical. We then test whether this stereotype translates into greater severity of punishment for women, relative to men, through a randomized, controlled laboratory experiment where the professional's role and ethical violation are held constant. The second study also tests whether the scope of our theory is correct by asking whether women are punished more severely for errors more broadly, or only for ethical violations, as we suggest. Finally, our third study tests whether the relationship between gender and punishment severity surfaces in a real-world field setting, using a sample of disciplinary hearings held to punish the ethical violations of attorneys in the U.S. In this field context, punishment decisions have important implications for individuals' careers. Our multi-method approach has significant advantages (Chatman & Flynn, 2005). By examining punishment severity in well-controlled laboratory experiments

designed to establish causality, we ensure that gender is the causal variable. By examining it in the field, we ensure that the phenomenon is relevant and has significant consequences in organizational settings.

Our research makes three important theoretical contributions. First, it documents disparate punishment for ethical violations as a novel form of gender discrimination. By doing so, it contributes to a body of research describing the forces that not only constrain but even *reverse* women's advancement at work (Berdahl, 2007; Brescoll et al., 2010; Ryan & Haslam, 2007). Although researchers have largely focused on how gender biases the perceptions of professionals (for reviews, see Cuddy, Glick, & Beninger, 2011 and Rudman & Phelan, 2008), less is known about the behavioral outcomes of these perceptions. Gender-biased perceptions can lead to unfair allocation of pay and promotions (Castilla & Bernard, 2010; Desai, Chugh, & Brief, 2014) and also to unethical behaviors designed to undermine women, such as sabotage (Rudman & Fairchild, 2004) and deception (Kray, Kennedy, & Van Zant, 2014). We find that discriminatory perceptions also manifest through punishments. Second, we find evidence for a new prescriptive stereotype. People expect women to be more ethical than men. This may help to explain why gender differences in ethics emerge (Franke, Crown, & Spake, 1997; Kennedy & Kray, 2013; Kray & Haselhuhn, 2012). Finally, we extend knowledge of factors affecting punishment severity (cf. Arvey & Ivancevich, 1980; Butterfield et al., 2005; Tetlock et al, 2007; Trevino, 1992; Wiltermuth & Flynn, 2013), an important issue at work. Our research is the first to document that gender predicts punishment severity.

Gender and the Prescription to be Ethical

It is well-known that women and men face different social expectations. To date, researchers have largely focused on stereotypes regarding warmth and competence (also known as communion and agency) (Conway, Pizzamiglio, & Mount, 1996; Williams & Best, 1982). Because women have traditionally held roles as homemakers rather than providers (Eagly, Wood, & Diekmann, 2000), women are expected to be more concerned with others and less assertive (Eagly & Karau, 2002; Heilman, 2001), as well as warmer and less competent than men (Eagly & Mladinic, 1989; Fiske, Cuddy, Glick, & Xu,

2002). These stereotypes have a prescriptive quality (Burgess & Borgida, 1999; Eagly & Karau, 2002; Heilman, 2001). That is, women are expected to behave in ways consistent with these stereotypes and face negative consequences when their behavior deviates from them.

We propose that ethicality is another characteristic prescribed more strongly for women than men. Ethicality can be distinguished from warmth. For instance, Fiske et al. (2002) noted that warmth involves some traits related to morality (e.g., sincerity), but also that their scales confound moral traits with sociality (e.g., good-naturedness, warmth). More recent research has distinguished morality and warmth (Brambilla & Leach, 2014; Goodwin, Piazza, & Rozin, 2014), reasoning that important moral traits such as honesty, fairness, and loyalty can be displayed without social warmth (e.g., friendliness and enthusiasm) (Goodwin, 2015). Existing theories of gender (Eagly et al., 2000; Fiske et al., 2002), have focused on warmth, not morality. Because warmth and morality are not synonymous, it is not obvious that women would be held to higher standards of morality. For instance, Prentice and Carranza (2002) found that the prescription to be principled was weaker for women relative to men. Goodwin et al. (2014, p. 166) said, “It remains an open question whether moral information, in general, is valued equally in men and women, relative to warmth information.”

We expect that women are held to a higher standard of ethicality in their professional performance than are men because of their historical association with communal roles (Eagly et al., 2000). Communal roles confer the responsibility of developing relationships with others and building community (Abele & Wojcizke, 2007, p. 754). Traits associated with morality, such as humility, commitment, and trustworthiness, play an important role in community-building, just as traits related to warmth, such as happiness, enthusiasm, and extroversion (Haidt, 2007). Prior theorists have suggested this possibility (Heilman, 2012), but it has not yet received empirical attention, though indirect evidence exists. For example, in their study of prescriptive gender stereotypes, Prentice and Carranza (2002) found that women faced intensified prescriptions to be spiritual and wholesome. Empirically, important ethical differences also emerge between women and men. Women are more prone to moral emotions of guilt and

shame (Cohen, Wolf, Panter, & Insko, 2011), they exhibit higher levels of moral character (Cohen, Panter, Turan, Morse, & Kim, 2014), more traditionally moral sociopolitical attitudes (Eagly, Diekmann, Johannesen-Schmidt, & Koenig, 2004), greater moral outrage at ethical compromises (Kennedy & Kray, 2013), higher and less flexible ethical standards (Franke et al., 1997; Kray & Haselhuhn, 2012), and lower moral disengagement (Detert, Trevino, & Sweitzer, 2008). Women are also less accepting of the use of deception (Dreber & Johannesson, 2008) and unethical negotiating tactics (Robinson, Lewicki, & Donahue, 2000) to secure monetary payoffs. These ethical differences by gender could reflect women's internalization of the higher ethical standards we expect them to face. On the basis of this logic, we propose the following hypothesis:

Hypothesis 1. *Women face stronger prescriptions for ethical behavior than men.*

If women face stronger prescriptions of ethicality than men, people should react more negatively to women who violate ethical standards compared to men. One important behavioral consequence could be imposing harsher punishment on women who violate ethical standards, as compared to men.

Punishing Ethical Violations

Punishment is the application of an aversive event or removal of a pleasing event in response to undesirable behavior (Arvey & Ivancevich, 1980; Ball & Sims, 1991). The desire to punish ethical violations is a widely held human motive; it reaffirms a community's ethical standards following a breach (Durkheim, 1925/1976), establishes norms for appropriate behavior (Feldman & March, 1981), and is intended to restore victims' dignity (Rucker, Polifroni, Tetlock, & Scott, 2004). The challenge for those in a punishment-seeking mindset is to act fairly (Tetlock, 2002; Tetlock et al, 2007).

Gender discrimination is likely to be visible in the context of punishment because it can be difficult to ascertain the level of punishment that is fair for a situation. What constitutes an ethical violation is often ambiguous (Wiltermuth & Flynn, 2013), and people tend to over-estimate the extent to which others share their ethical views (Flynn & Wiltermuth, 2010). Because there is little consensus regarding ethical standards or how violations should be punished, the context is ripe for gender

discrimination, which tends to emerge when decision-makers are afforded substantial discretion under a guise of objectivity (Castilla, 2008; Castilla & Benard, 2010; Uhlmann & Cohen, 2005). Under these circumstances, gender may subconsciously influence evaluators' perceptions, leading to discrepant outcomes that tend to systematically disfavor women (Gorman, 2005; Reskin, 2000).

Seemingly irrelevant factors influence the severity of punishment (Ball & Sims, 1991), including the power (Mooijman, van Dijk, Ellemers, & van Dijk, 2015; Wiltermuth & Flynn, 2013) and political ideology of the punisher (Tetlock et al., 2007), the ambiguity of the violation (Tetlock et al., 2007), and a sense of a threatened social order (Rucker et al., 2004). While research has not examined gender as a predictor of punishment, there are reasons to think women could be punished *less* (not more) severely than men for ethical violations. People generally view women positively (Glick & Fiske, 1996) and see them as well-intentioned (Fiske et al., 2002), so they may be inclined to give women's intentions the benefit of a doubt or make external attributions for behavior that is at odds with benevolent stereotypes of women's intentions. Moreover, higher-status perpetrators are punished more harshly (Fragale et al., 2009; Tetlock et al., 2007), and women are often considered to hold lower status in society than men (Ridgeway, 2001).

Nevertheless, because ethical behavior is more strongly prescribed for women than men, we predict that women will be punished more severely than men following an ethical violation at work. Violations of gender stereotypes are met with punishment (for a review, see Rudman & Phelan, 2008), and more severe punishment is given for more severe violations. Severity is subjective, however. To the extent women are held to higher ethical standards than men, observers may perceive women's ethical violations to be more severe than men's because the higher standard women are held to makes even minor violations noticeable and worthy of condemnation. This pattern has been observed in other contexts, such as negotiation. When women and men exhibit identical assertiveness, women incur greater dislike and rejection than do men because people expect less assertiveness from women (Amanatullah & Tinsley, 2013). Existing empirical evidence provides preliminary support for this hypothesis regarding punishment

severity. Some research suggests women are punished more severely than men for violating norms through whistle-blowing (Rehg, Miceli, Near, & Van Scotter, 2008) or workplace deviance (Bowles & Gelfand, 2010). However, these studies do not directly focus on ethical violations and leave room for a deeper exploration of the issue. Whistle-blowing is of ambiguous ethicality (Waytz, Dungan, & Young, 2013) and perceptions of retaliation were self-reported in Rehg et al.'s (2008) data. Moreover, only one of Bowles and Gelfand's (2010) studies focused on gender, and it left the psychological mechanism unspecified. Here, we build on their findings to explore gender disparities in punishment in greater depth.

Hypothesis 2. *Women will be punished more severely than men for committing an ethical violation in their profession.*

To the extent that women receive harsher punishments than men for the same ethical violations, this is an important issue in organizations. Punishment equity requires that similar punishments are assigned for similar violations (Butterfield et al., 2005; Trevino, 1992). Harsh punishment leads to adverse effects like anxiety, aggression, withdrawal, and sabotage (Arvey & Ivancevich, 1980). Women could therefore suffer these adverse side effects more than men if punished in a discriminatory manner.

Gender Composition of the Decision-Making Group as a Moderating Factor

One outstanding question is whether the gender disparity in punishment is affected by the representation of women among key decision-makers (cf. Duguid, 2011; Duguid, Lloyd, & Tolbert, 2012). Prior sociological work suggests that the gender composition of decision-makers serves as an important moderating factor for the level of discrimination present in work outcomes. When the people who control work outcomes are more demographically similar to those whom their decisions affect, discrimination diminishes (Gorman, 2005; Kulis, 1997; Konrad and Pfeffer, 1991; Pfeffer and Davis-Blake, 1987). For instance, Ely (1995) found evidence of stronger stereotyping in firms where women were under-represented in positions of power, and suggested that the presence of women might diminish stereotyping by making gender a less salient category. Drawing from this research, we predict that the

gender disparity in punishment severity is smaller when women are more highly represented among the decision-makers allotting punishment.

Hypothesis 3. *The gender of the decision-maker(s) responsible for allotting punishment moderates the extent to which women are punished more severely than men for similar ethical violations.*

OVERVIEW OF STUDIES

We conducted three studies to test our hypotheses. Study 1 investigates the relation between gender and the prescriptive stereotype to be ethical. Study 2 explores punishment severity by manipulating a professional's gender in a hypothetical scenario, while holding other factors constant. The experimental design establishes that the gender of the individual who commits the ethical violation, rather than other factors, is the causal factor that drives differential punishment severity. Study 2 also compares punishment for ethical violations to punishment for accidental errors. By doing so, it tests our theoretical mechanism: To the extent that women are punished more severely than men due to a prescriptive stereotype to be ethical, we would expect harsher punishment only for ethical violations, but not for accidental errors. Study 3 tests whether the relationship between gender and punishment severity surfaces in a real-world field setting, where punishment decisions have important career implications. Additionally, Study 3 tests a key moderating factor: the representation of women among those assigning punishment.

STUDY 1: DOES GENDER BIAS ETHICAL PRESCRIPTIONS?

Study 1 explored the prescription to be ethical using an experiment with a sample of working adults. We tested our first hypothesis that ethical behavior is more strongly prescribed for women than for men. Here and subsequently, we conducted our experiments in line with current recommendations for experimental research (Simmons, Nelson, & Simonsohn, 2011). We pre-determined our sample sizes and disclose all of the dependent variables that were measured.

Methods

Sample. Participants ($N = 202$) were recruited from an online pool. The sample was 45% women and had a mean age of 34.46 ($SD = 10.48$) years. In prior studies recruited from this pool, we found participants to have 11.36 ($SD = 9.46$) years of work experience, on average, with most (51%) employed full-time, some (24%) part-time, and 25% not currently employed outside the work they do on the web site.

Design and procedure. The study used a 2-condition (Attorney's gender: Male, Female), between-participants design. First, participants read the following scenario describing an attorney. The scenario was designed to provide enough background information on the attorney for participants to feel they were making informed judgments, and to distract participants from gender as the topic of our investigation. Only the attorney's gender (as indicated by first name and pronouns) varied between conditions:

Imagine that you know a lawyer, Jennifer / John Thompson. She / he studied Political Science as an undergrad and went to Villanova University for law school. She / He is relatively young, but has a strong track record. You have only met Jennifer / John twice to date. You know that she / he is interested in art and wine, and enjoys running on the weekend. Jennifer / John is a litigator. In court, she / he mostly serves executives at business firms. For instance, Jennifer / John is currently defending a businessperson who served for the last 7 years as a key executive at EnviroHeat, a company that manufactures wood pellets, the material burned in pellet stoves to generate heat. Pellet stoves are environmentally friendly and are primarily used to warm homes. Jennifer's / John's client took the executive job as a favor to some investors. The market for pellet stoves has been brutal with oil prices so low. For 7 years, the executive worked hard for EnviroHeat, but the company had market-driven (not company-specific) problems. When oil prices dropped severely last year, company earnings declined further, and EnviroHeat's investors fired the executive and refused to pay severance. Investors claimed the executive was fired "for cause." Investors have some evidence that the executive spent too much time golfing during work hours, and tended to hire friends into important executive roles. Jennifer / John is suing the investors for her / his client to get the severance pay.

Then, participants took a moment to imagine the attorney, Jennifer or John, and rated the degree to which they expected her/him to display a number of traits. We included traits measuring ethicality (e.g., loyal, trustworthy, has integrity), warmth (e.g., warm, agreeable), and competence (e.g., competent, capable).

Ethical prescriptions measure. To measure ethical prescriptions, we asked participants to rate the extent to which they expected the lawyer to display 16 moral traits, using a scale from 1 (not at all) to 9 (very much). Nine traits (loyal, trustworthy, honest, principled, fair-minded, responsible, courageous,

dependable, and sincere) were drawn from Goodwin's (2015; Goodwin et al., 2014) list of high-morality, lower-warmth traits. We selected these traits in order to ensure that the study tested for a stereotype related to ethicality, not warmth. To those nine traits, we added three (ethical, has integrity, has good values) designed to increase face validity, and four (ruthless, immoral, selfish, greedy) that were reverse-scored. Together, the items formed a reliable scale ($\alpha = .96$). Using the same 9-point scale, participants also reported their expectations of the lawyer's warmth and competence, using items from prior research (Fiske et al., 2002). To measure warmth, we used the terms enthusiastic, warm, sociable, and agreeable ($\alpha = .84$). To measure competence, we used the terms confident, competent, intelligent, and capable ($\alpha = .93$).

Results

To ensure that each trait scale measured a distinct construct, we first conducted a factor analysis on the 24 items. Using direct oblimin rotation, four distinct factors with eigenvalues greater than one emerged, explaining 75.97% of the variance. The factors corresponded to: positive ethicality (12 items, eigenvalue = 12.53, 52.20% of the variance), negative ethicality (4 items, eigenvalue = 3.49, 14.53% of the variance), warmth (4 items, eigenvalue = 1.98, 8.24% of the variance), and competence (4 items, eigenvalue = 1.23, 5.11% of the variance). Consistent with previous research, this analysis confirmed that ethicality and warmth are distinct constructs. For subsequent analyses, we combined positive and negative ethicality (reverse-scored) into one measure ($\alpha = .96$).

We next conducted a repeated-measures ANOVA with the three trait scales (ethicality, warmth, and competence) as a within-participants factor. We included the attorney's gender and participant's gender as between-participant factors. Three main effects and an interaction emerged. First, a main effect for trait showed that participants perceived the attorney to be more competent ($M = 7.78$, $SE = .08$) than ethical ($M = 6.40$, $SE = .10$, $p < .001$) or warm ($M = 6.55$, $SE = .09$), $F(2, 197) = 130.74$, $p < .001$, $\eta_p^2 = .57$. Ratings of ethicality and warmth were not significantly different, $p = .10$. Second, a main effect of participant gender showed that women ($M = 7.08$, $SE = .11$) perceived the attorney more positively across the three dimensions than did men ($M = 6.75$, $SE = .10$), $F(1, 198) = 4.60$, $p = .03$, $\eta_p^2 = .02$. Finally, the

main effect for attorney's gender ($F(1, 198) = 3.92, p = .049, \eta_p^2 = .02$), was qualified by a trait X attorney's gender interaction, $F(2, 197) = 6.50, p = .002, \eta_p^2 = .06$. The interaction suggests that the attorney's gender impacted perceptions of him/her in terms of some traits, but not others.

To understand this interaction, we analyzed the three trait scales separately. Three separate ANOVAs included attorney's and participant's gender as between-participant factors. First, we explored our central research question: Was the attorney expected to be more ethical when she was female rather than male? A main effect of attorney's gender supported Hypothesis 1, $F(1, 198) = 12.35, p = .001, \eta_p^2 = .06$. The female attorney was expected to be more ethical ($M = 6.77, SE = .15$) than the male ($M = 6.04, SE = .15$). The pattern of means suggests that the attorney was expected to be ethical regardless of gender, as both means were above the scale mid-point, but this expectation was intensified the female lawyer. No other main effects or interactions emerged ($ps > .48$).

The attorney's gender did not impact warmth or competence stereotypes, perhaps because of the vivid description of the attorney's role that was uniformly provided. For warmth, only a main effect of participant gender emerged. Women ($M = 6.77, SE = .14$) expected the attorney to be warmer than did men ($M = 6.34, SD = .12$), $F(1, 198) = 5.38, p = .02, \eta_p^2 = .03$. No other main effects or interactions were significant ($ps > .33$). For competence, again, only a main effect of participant gender emerged. Women ($M = 7.99, SE = .12$) expected the attorney to be more competent than did men ($M = 7.57, SD = .11$), $F(1, 198) = 6.15, p = .01, \eta_p^2 = .03$. Other main effects and interactions were non-significant ($ps > .83$).

The results support Hypothesis 1. Women face intensified ethical prescriptions, even when in an identical professional role. Ethical characteristics are more strongly prescribed for women than for men.

STUDY 2: PUNISHMENT SEVERITY IN A LABORATORY EXPERIMENT

In Study 2, we aimed to test whether gender affects how severely a professional is punished for an identical ethical violation. To do so, we used an experimental design. In addition, Study 2 tests a key boundary condition: whether women professionals are punished more for accidental mistakes or only for intentional ethical violations. We focus on ethical violations in light of the content of prescriptive gender

stereotypes. Because women face relaxed prescriptions surrounding competence-related qualities (Biernat & Kobrynowicz, 1997; Prentice & Carranza, 2002), we do not expect them to be punished more harshly than men for accidental mistakes at work. Thus, this moderating factor provides an additional test of our theoretical mechanism. If women face intensified ethical prescriptions, they should face more severe punishment than men for ethical violations but not accidental mistakes.

Sample

Master's of Business Administration students ($N = 194$) at two business schools completed a survey about work outcomes. Women comprised 35% of the sample. The average age was 26.55 years ($SD = 3.76$). The sample was 5% African American, 17% Asian, 74% Caucasian, 4% Latino/Hispanic, and 1% who selected "other" backgrounds.

Design and Procedure

The study employed a 2 (Target's gender: Woman, Man) X 2 (Intentionality: Accidental mistake, Ethical violation) between-participants design. Our expectation, which we confirm through a pre-test, was that only intentional violations are construed as ethical violations; in contrast, accidental mistakes are considered non-ethical violations. Participants were randomly assigned to one of the four conditions. They read a scenario about a hospital manager who filed an erroneous Medicare claim requesting larger payment for medical services than was actually owed. The error had serious negative consequences for the organization: the Medicare office discovered the error, accused the hospital of fraud, and assessed a heavy fine of \$100,000. We manipulated the gender of the hospital manager by changing the name and pronouns used to refer to the manager. We also manipulated whether the error was an ethical violation by stating that it was made accidentally or intentionally. The scenario's full text read:

(Jake / Jill) Moranty, a 40-year-old single (man / woman), is employed as the lead manager of Bethany Central Hospital in Kansas City, Missouri. Moranty has worked at the hospital for twenty years. One of Moranty's employment duties is to fill out the monthly Medicare reimbursable claims report for all surgeries performed in the emergency ward. (Mr. / Ms.) Moranty fills out the paperwork and then, once the Medicare office receives the report, it pays the hospital the amount owed for the reported services. This month, Moranty filed an erroneous claims report. In it, (he / she) (accidentally / intentionally) reported several of the procedures twice, ultimately asking for \$10,000 more than what the hospital was owed. The Medicare office

caught the discrepancy in the figures and has accused the hospital of Medicare fraud. The event has triggered an enormous amount of bad press for the hospital, and the hospital ultimately had to pay \$100,000 to the federal government to settle the charges.

This scenario has two key strengths. It describes an ethical violation that is non-relational (Dawson, 1992; Franke et al., 1997) and pro-organizational (Vadera & Pratt, 2013). Ethical violations in the context of close relationships (Dawson, 1992) or that benefit oneself (rather than the organization) (Amanatullah & Morris, 2010) could violate gender prescriptions of warmth. By examining a non-relational, pro-organizational ethical violation, we avoid this confound and provide a conservative test of our hypotheses. After reading this scenario, participants answered a series of questions measuring how severely they believed the hospital manager should be punished.

Punishment severity measure. To provide a measure of punishment severity, participants reported the most appropriate jail sentence for Moranty, using a scale from 0 to 500 days. This measure captured willingness to mete out a severe sentence.¹

Results

Pre-test results. A pre-test with 103 respondents confirmed that the intentional error was perceived as an ethical violation, more so than the accidental mistake. Using a 7-point scale, respondents indicated their agreement that Moranty committed and ethical violation, acted unethically, was morally wrong, and that the incident had nothing to do with Moranty's ethical standards (reverse-scored) ($\alpha = .96$). The intentional error was perceived as an ethical violation ($M = 5.83$, $SD = 1.26$) to a greater degree than the accidental mistake ($M = 2.78$, $SD = 1.34$), $F = 141.03$, $p < .001$, $\eta_p^2 = .59$. Professional's gender ($F = 1.01$, $p = .32$, $\eta_p^2 = .01$) and the interaction term ($F = 0.77$, $p = .38$, $\eta_p^2 = .01$) were non-significant.

Our first analyses included a factor for the business school where data were collected. One main effect emerged. Students at one business school showed greater punishment severity toward the hospital

¹ On an exploratory basis, we also measured participants' punitive mindset ($\alpha = .85$) (Tetlock, 2002; Tetlock et al., 2007) and their ratings of the professional's social status ($\alpha = .62$). For punitive mindset, only a main effect of intentionality emerged, $F(1, 189) = 146.87$, $p < .001$, $\eta_p^2 = .44$ (other $ps > .10$). For social status, no statistically significant effects emerged ($ps > .12$).

manager. However, no significant interactions emerged so we collapsed across the samples. Analyses including a covariate controlling for the sample yield virtually identical results to those shown below.

Primary analyses. We analyzed the data using ANOVA with target's gender and the intentionality of the error as between-participant factors. First, we asked: Did the target's gender affect punishment severity for errors in general, or only for ethical violations, as we predicted? For punishment severity, one main effect and an interaction emerged. A main effect emerged for intentionality, $F(1, 188) = 30.43, p < .001, \eta_p^2 = .14$. More severe punishments were recommended for ethical violations ($M = 107.04$ days, $SD = 141.83$) than for accidental mistakes ($M = 21.28$ days, $SD = 68.10$). The main effect for target's gender was non-significant, $F(1, 188) = 0.76, p = .39, \eta_p^2 = .004$. As expected, there was also an interaction between target's gender and intentionality, $F(1, 188) = 4.91, p = .03, \eta_p^2 = .03$.

To explore the interaction, we examined gender differences in punishment severity within each intentionality condition. The manager was punished more severely for an ethical violation relative to an accidental mistake regardless of whether the manager was male ($t[97] = 2.76, p = .007, d = 0.56$) or female ($t[92] = 4.88, p < .001, d = 0.98$). For an accidental mistake, no difference in punishment severity emerged for the female versus male manager, $t[97] = -1.27, p = .21, d = 0.25$. However, for an intentional ethical violation, the female manager was punished more severely than the male manager, at a marginally significant level, $t[92] = 1.74, p = .09, d = 0.56$. Means are shown in Figure 1.

These results provide preliminary support for Hypothesis 2. Following an ethical violation, a manager was punished more severely when a woman rather than a man. No difference by target's gender emerged for an accidental mistake. These findings provide additional support for the notion that women face an intensified prescription to be ethical. In Study 3, we sought to replicate this finding using a larger, field-based sample. The larger sample size also allows us to examine Hypothesis 3.

STUDY 3: PUNISHMENT SEVERITY IN ATTORNEY DISCIPLINARY HEARINGS

Study 3 extended knowledge by examining punishment severity in a field setting. To probe the external validity of our findings in Study 2, we created an original database of archival decisions rendered

in attorney disciplinary hearings. We examined the severity of punishments applied to attorneys who had violated the American Bar Association's (ABA) rules of professional conduct. Attorneys who are found to have violated an ethical rule receive punishments of differing severity through these hearings, with important implications for their subsequent careers and professional status. This database also allowed us to explore how the representation of women among judges in the disciplinary hearings impacts punishment severity.

For many reasons, this setting provides an ideal test of our theory. First, punishment in this setting is systematically documented, as the outcomes of attorney disciplinary hearings are reported in publically available reports of court outcomes. Punishments meted out to professionals can take the form of lower year-end bonuses, assignment to less desirable teams or projects, slower promotions, or removal from leadership positions. These punishments are often private and rarely documented systematically, but this data set offered a comprehensive record. Second, what constitutes an ethical violation is unusually clear in this setting. The ABA's ethical rules for attorneys are carefully defined and attorneys are familiar with them. In fact, admission to a state's bar requires attorneys to pass an ethics exam that tests their knowledge of the ABA's Model Rules of Professional Conduct (MRPC). Moreover, reports of disciplinary hearings explicitly document each rule that the attorney is accused of committing, allowing for a controlled comparison of punishment across different types of ethical violations. Third, punishment severity in these cases is indexed in a straightforward way. In other settings, the severity of one type of punishment relative to another is often ambiguous. Whether a lower bonus or a foregone promotion is more severe is open to interpretation. In this setting, however, punishment severity has been thoughtfully delineated by the ABA. Finally, the panel of judges overseeing these hearings receives explicit guidance regarding how the rules should be applied and how to consider aggravating and mitigating factors in the setting of punishments. The context thus provides a conservative test of our hypothesis. If women are punished more severely in this context, even greater differentiation may be found in other contexts, where judgment is less public, more intuitive and automatic, and parties may feel freer to act on biases.

Sample

We created a database of all states' attorney disciplinary cases heard by state-level appellate courts in the year 2008, as published in Westlaw's ALLSTATES database of legal cases. Westlaw is a comprehensive legal resource that includes a searchable database of decided cases. From among these cases, we selected all of those cases that analyzed the alleged ethical violations using the current version of the MRPC, which made it possible to uniformly control for which ethical rules were violated in each case. The MRPC, written and promulgated by the ABA in 1983, codifies the ethical obligations of the legal profession. Although states are under no express obligation to adopt the MRPC, thirty-eight states had adopted the most recent version in full by 2008. Our final sample of cases applying the most recent version of the MRPC includes a total of 482 cases heard across 33 states. The states with the most cases in the sample are New Jersey (111 cases), Colorado (59 cases), Indiana and Kansas (each with 35 cases).

Critical to our immediate inquiry, the MRPC includes a set of uniform standards to govern the imposition of punitive sanctions for attorneys who violate the rules. These standards are intended to promote consistency in the application of punishment for ethical transgressions within and among separate jurisdictions. Here, the ABA provides a list of the possible sanctions for violations of the rules of professional conduct, as well as general standards for when each is appropriate. Possible sanctions include, in increasing severity: admonition (or private reprimand),² public reprimand (declaration of the impropriety of the lawyer's action, without interrupting the right to practice), probation (a period of increased scrutiny, without interrupting the right to practice), suspension (a temporary interruption in the lawyer's right to practice), and disbarment (extinguishing the transgressor's status as a lawyer).

Measures

Punishment severity. Our dependent variable, *punishment severity*, is an ordinal variable. It reflects different categories of disciplinary outcomes using levels specified by the ABA. Specifically, the variable is coded "0" if the attorney was publically reprimanded or censured, "1" if the attorney was put

² Because admonitions or private reprimands are not a matter of public record, this category of sanction is not represented in our sample of cases.

on probation, “2” if the attorney was suspended for a discrete amount of time, “3” if the attorney was indefinitely suspended, and “4” if the attorney was disbarred. In cases where more than one category of punishment was employed, we coded this variable as the more severe punishment. For example, if an attorney was suspended for six months to be followed by a year of probation, we coded their punishment as a suspension for a discrete amount of time.

Attorney’s gender. The attorney’s gender was coded from the text of the disciplinary hearing summaries as reported in Westlaw based on the name and pronoun (he/she) used to refer to the accused attorney. We coded the gender variable “0” for men and “1” for women. Our sample included 83 cases against female attorneys, or 17% of the full sample.

Percentage of women on judges’ panel. Each hearing was overseen by a panel of judges. To provide a test of Hypothesis 3, we include the percentage of women sitting on the judges’ panel. The list of judges involved in the hearing was reported in over half of our sample of cases. We identified each judge’s gender by matching their names to their pictures and biographies on court websites.

Control variables. We control for a number of factors that influence punishment severity and might act as third variables to create a spurious relationship between gender and punishment severity. To control for the type of ethical violations committed in each case, we include fixed effects for each individual rule –as codified in the MPRC-- that was found to be violated. Our full sample of cases involves 39 discrete rules were violated. Each rule is coded “1” if it was violated in a specific case and “0” if it was not. Those violated most often include Rule 8.4, Misconduct (n= 313); Rule 1.3, Diligence (n= 186); Rule 1.4, Communication (n= 181); Rule 1.15, Safekeeping Property (n= 141); and Rule 8.1, Bar Admission and Disciplinary Matters (n= 114). To avoid overestimation of our models, we only include controls for each rule that had at least 15 observations, thereby appearing in at least 3% of the full sample of cases. For each case, we also included a control for the aggregated *total number of rules violated* in the case. This count variable ranged from a low of 1 to a high of 13 rule violations.

To account for the possibility that the number of decision-makers involved in administering a punishment may affect the severity of the punishment rendered, we include a control for the *total number of judges on the hearing panel*, as reported in each case report.

Classical punitive theory in sociology suggests that the severity of punishment for an offense is likely to be systematically related to the economic value of the accused's professional contributions to society (Rusche & Kirchheimer, 1936). According to this perspective, attorneys may be less severely punished when they practice in states where there are fewer attorneys, making their services more valuable to the public. To account for this possibility, we include a variable capturing the general economic market in which disparate states' legal labor systems are embedded by controlling for the *per capita legal saturation*. This variable reflects the number of lawyers per 10,000 residents within the US state in which each disciplinary hearing was held (DC inclusive). This data was compiled for 2008 and made publically available by AveryIndex. Descriptive statistics and correlations are shown in Table 1.

Statistical Methods

Because our dependent variable, *punishment severity*, is coded as an ordinal variable, we employ ordered logistic regression using the *ologit* command in Stata 14. This model essentially allows for the application of the standard logit model – which is applied to dichotomous outcome variables – to dependent variables that consist of more than two categories that have a natural underlying order to them. As a robustness check, we replicate our results with a simple probit model predicting the likelihood of the highest category of punishment: disbarment. In these models, the dependent variable is coded “1” if the accused was disbarred, and “0” if the accused received any lesser form of punishment.

To control for potential heteroskedasticity, we employed robust standard errors across all models, clustered by the state in which the case took place (e.g., Nordgren & McDonnell, 2011). Clustering standard errors in this manner also takes into account the likelihood that state courts will afford greater precedential value to their own past cases than to cases held in other states. If so, disciplinary hearing outcomes will be correlated more closely within-state than between-states.

Results

Results are shown in Table 2. Model 1 includes all 482 cases in the full sample. Models 2-5 include controls for the number of judges sitting on the panel and the percentage of female judges on the panel. Because only a subset of the cases in the larger sample disclosed the judges empaneled for the hearing, the number of observations in Models 2-5 ($n = 270$) is lower than that in Model 1. Models 3 and 5 include the interaction of the disciplined attorney's gender with the percentage of female judges on the hearing panel, testing Hypothesis 3.

Replicating the same pattern of findings observed in Study 2, in Study 3 we find robust support across all models for our hypothesis that women are punished more severely than men. In Models 1-3, we find a significant positive association between gender and punishment severity. This finding suggests that, controlling for the ethical violation committed, female attorneys in these cases are generally assigned more severe punishments for their offense, relative to their male counterparts. This finding is replicated in the probit models depicted in Models 4-5, which suggests that women are more likely to be disbarred than men, controlling for the precise ethical violation committed. Post-estimation margins analysis of Model 4 with all control variables held to their means indicates that females have a 106% higher likelihood of being disbarred than males (at .35 and .17, respectively).

We also find strong support for Hypothesis 3, as the interaction of the disciplined attorney's gender with the percentage of female judges on the hearing panel is both negative and significant in both models. This indicates that the extent of discriminatory punishment is mitigated when women are more highly represented among the disciplinary hearing decision-makers. To assist with the interpretation of this finding, we plot the predicted values derived from post-estimation margins analysis of Model 3 in Figure 2. As the figure shows, when the percentage of females on the judges' panel was 1 standard deviation below the mean, a female attorney's likelihood of being disbarred is more than twice that of a male attorney. However, the likelihood of being disbarred is nearly identical between males and females when the percentage of women on the judges' panel is one standard deviation above the mean. This result

suggests that greater representation of women among decision-makers who allot punishment can temper disparate punishments allotted to male and female professionals.

GENERAL DISCUSSION

It is vital for punishment to be distributed equitably based on the severity of the actions committed (Trevino, 1992). Yet, in the current research, we find evidence that the gender of the actor committing the ethical violation matters. Specifically, women face intensified prescriptions to be ethical and they tend to incur more severe punishments than men following ethical violations. These findings have important theoretical and practical implications.

Theoretical Contributions

Our research documents a new barrier faced by women in organizations. It is well-known that professional women face steeper social penalties for behaving in ways that benefit their careers, such as participating actively (Brescoll, 2011), displaying high levels of performance (Rudman & Fairchild, 2004), and negotiating (Bowles, Babcock, & Lai, 2007). We break new ground by documenting that women are also held to a higher *ethical* standard as they strive to excel in their careers, and suffer disproportionate punishment when they violate this prescription.

The ethical double standard is especially problematic because women are likely to think and act similarly to men when they hold the same occupational roles (Eagly et al., 2000). In Franke et al.'s (1997) meta-analysis, men's and women's ethical perceptions showed 84% overlap, and ethical differences diminished in samples with greater work experience. These trends spell trouble for professional women.

Punishment severity may be one mechanism by which women lose professional status more easily than men after achieving it. People appoint women to more precarious leadership roles (Ryan & Haslam, 2007) and accord them less status than men following mistakes (Brescoll et al., 2010). They also target professional women with incivility, harassment, and deception (Berdahl, 2007; Cortina, 2008; Kray et al., 2014). In light of this mistreatment, it is perhaps not surprising that women's professional

satisfaction is lower than men's, even among highly qualified women who have ostensibly had excellent opportunities for advancement (Ely et al., 2014).

Our findings could also imply that downward spirals are more likely to emerge within the careers of women. When people are punished unjustly, their citizenship behavior and performance decline (Cohen-Charash & Spector, 2001), precipitating additional status losses and further punishment, possibly culminating in their leaving the workplace. People who feel treated unfairly become disaffected and are more likely to protest, resign, and cheat (Tyler, 1990). Although we cannot speak to whether women feel treated unjustly following ethical violations, we show that they have reason to do so.

Our research also helps to explain why gender differences in ethical attitudes and behaviors emerge (Franke et al., 1997; Gilligan, 1982; Kennedy & Kray, 2013; Kray & Haselhuhn, 2012): People believe that women should be more ethical than men. Our work suggests women receive this message not only during childhood socialization but also within their work organizations (cf. Ely & Padavic, 2007).

Practical Contributions

Our research suggests that organizations need to establish more systematic approaches to punishing ethical violations in order to avoid unfairly disadvantaging women. When punishments are left to managerial discretion, they are not applied equitably. Instead, punishment is biased by the gender of the person who has committed an ethical violation, to the detriment of women's careers. The fact that this trend emerged in attorney disciplinary cases, where ethical rules are relatively codified and the process of deliberation is unusually structured, highlights the robustness of the effect.

Although we believe responsibility for equitable punishment lies with those who mete it out, our findings do suggest that professional women should take a conservative approach to ethical decision-making. Inferences drawn about ethical standards from norms set by male colleagues are not necessarily useful for women, as higher standards are prescribed for them.

In addition, our research suggests that it may be especially important for women to refute false accusations of unethical behavior. Apologies convey empathy (Brooks, Dai, & Schweitzer, 2014), and

may therefore be intuitively appealing to women who seek to conform to warmth prescriptions. However, because ethical violations are viewed more negatively for women than men, apologies may be especially damaging and ineffective as compared to denials of wrongdoing (Kim, Ferrin, Cooper, & Dirks, 2004).

Future Directions

Future research could explore correctives for the bias we document. Tetlock et al.'s (2007) fair-but-biased-yet-correctible (FBC) model holds that people try to correct their judgment when they perceive that it strays from their own private standards of good judgment. Discriminatory punishment is likely to be unintentional. Therefore, transparent decision-making processes with accountability for outcomes (Castilla, 2008, 2011; Castilla & Benard, 2010; Lerner & Tetlock, 1999; Tetlock & Mitchell, 2009) may enable those responsible for setting punishments to detect and correct their gender-related biases.

Future research could also examine how gender affects the fairness of other dimensions of punishment. For instance, Butterfield et al. (2005) summarized the importance of contingency (i.e., association with particular behavior), constructiveness, causal accounts, timeliness, and privacy for equitable punishment. It is possible that women suffer not only more severe punishments than men, but also worse treatment along these other dimensions. To the extent that women suffer more diffuse, unconstructive, unjustified, untimely, or public criticism, it could make sense that they respond more negatively and emotionally, as recently suggested by Nobel Laureate Sir Tim Hunt (Bever, 2015).

Conclusion

Our investigation breaks new ground by describing an important, but under-studied form of discrimination: more severe punishments for women than men, following an ethical violation. To help explain this phenomenon, we also document a new prescriptive gender stereotype: Women are expected to be more ethical than men, even when they hold an identical professional role. Our findings have important implications for women's careers. It may be necessary for managers to more closely monitor punishments for gender discrimination if they wish to ensure equal opportunity in their organization.

REFERENCES

- Abele, A. E., & Wojciszke, B. 2007. Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, 93: 751–763.
- Amanatullah, E. T., & Morris, M. W. 2010. Negotiating gender roles: Gender differences in assertive negotiating are mediated by women's fear of backlash and attenuated when negotiating on behalf of others. *Journal of Personality and Social Psychology*, 98: 256-267.
- Amanatullah, E. T., & Tinsley, C. H. 2013. Punishing female negotiators for asserting too much...or not enough: Exploring why advocacy moderates backlash against assertive female negotiators. *Organizational Behavior and Human Decision Processes*, 120: 110-122.
- Arvey, R. D., & Ivancevich, J. M. 1980. Punishment in organizations: A review, propositions, and research suggestions. *Academy of Management Review*, 5: 123–132.
- Ball, G. A., & Sims, H. P. 1991. A conceptual analysis of cognition and affect in organizational punishment. *Human Resource Management Review*, 1: 227–243.
- Beckman, C. M. & Phillips, D. J. 2005. Interorganizational determinants of promotion: Client leadership and the attainment of women attorneys. *American Sociological Review*, 70: 678-701.
- Belliveau, M. A. 2005. Blind ambition? The effects of social networks and institutional sex composition on the job search outcomes of elite coeducational and women's college graduates. *Organization Science*, 16: 134-150.
- Belliveau, M. A. 2012. Engendering inequity? How social accounts create vs. merely explain unfavorable pay outcomes for women. *Organization Science*, 23: 1154-1174.
- Berdahl, J. L. 2007. The sexual harassment of uppity women. *Journal of Applied Psychology*, 92: 425-437.
- Bever, L. Nobel prize-winning scientist Tim Hunt resigns after commenting on “the trouble with girls.” *Washington Post*. Retrieved from <https://www.washingtonpost.com/news/morning-mix/wp/2015/06/11/nobel-prize-winning-scientist-tim-hunt-resigns-position-after-commenting-on-the-trouble-with-girls/>
- Biernat, M., & Kobrynowicz, D. 1997. Gender- and race-based standards of competence: Lower minimum

- standards but higher ability standards for devalued groups. *Journal of Personality and Social Psychology*, 72: 544-557.
- 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.
- Bowles, H. R., & Gelfand, M. 2010. Status and the evaluation of workplace deviance. *Psychological Science*, 2: 49-54.
- Bowles, H. R., & McGinn, K. L. 2008. Untapped potential in the study of negotiation and gender inequality in organizations. In J. P. Walsh & A. P. Brief (Eds.) *Academy of Management Annals*, 2: 99-132.
- Brambilla, M., & Leach, C. W. 2014. On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition*, 32: 397-408.
- Brescoll, V. L., Dawson, E., & Uhlmann, E. L. 2010. Hard won and easily lost: The fragile status of leaders in gender-stereotype-incongruent occupations. *Psychological Science*, 1640-1642.
- Brescoll, V. L. 2011. Who takes the floor and why: Gender, power, and volubility in organizations. *Administrative Science Quarterly*, 56: 622-641.
- Brooks, A. W., Dai, H., & Schweitzer, M. E. 2014. I'm sorry about the rain! Superfluous apologies demonstrate empathic concern and increase trust. *Social Psychological and Personality Science*, 5: 467-474.
- Burgess, D. & Borgida, E. 1999. Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy, and Law*, 5: 1-28.
- Butterfield, K. D., Treviño, L. K., Wade, K. J., & Ball, G. A. 2005. Organizational punishment from the manager's perspective: An exploratory study. *Journal of Managerial Issues*, 17: 363-382.
- Castilla, E. J. 2008. Gender, race, and meritocracy in organizational careers. *American Journal of Sociology*, 113: 1479-1526.
- Castilla, E. J., & Benard, S. 2010. The paradox of meritocracy in organizations. *Administrative Science*

Quarterly, 55: 543-676.

Castilla, E. J. 2011. Bringing managers back in: Managerial influences on workplace inequality.

American Sociological Review, 76: 667-694.

Chatman, J. A., & Flynn, F. J. 2005. Full-cycle micro-organizational behavior research. *Organization*

Science, 16: 434-447.

Cohen-Charash, Y., & Spector, P. E. 2001. The role of justice in organizations: A meta-analysis.

Organizational Behavior and Human Decision Processes, 86: 278-321.

Cohen, T. R. Panter, A. T., Turan, N., Morse, L., & Kim, Y. 2014. Moral character in the workplace.

Journal of Personality and Social Psychology, 107: 943-963.

Cohen, T. R., Wolf, S. T., Panter, A. T., & Insko, C. A. 2011. Introducing the GASP scale: A new

measure of guilt and shame proneness. *Journal of Personality and Social Psychology*, 100: 947-966.

Conway, M., Pizzamiglio, M. T., & Mount, L. 1996. Status, communality, and agency: Implications for

stereotypes of gender and other groups. *Journal of Personality and Social Psychology*, 71: 25-38.

Cortina, L. M. 2008. Unseen injustice: Incivility as modern discrimination in organizations. *Academy of*

Management Review, 33: 55-75.

Cuddy, A. J. C., Glick, P., & Beninger, A. 2011. The dynamics of warmth and competence judgments,

and their outcomes in organizations. *Research in Organizational Behavior*, 31: 73-98.

Dawson, L. M. 1992. Will feminization change the ethics of the sales profession? *Journal of Personal*

Selling & Sales Management, 13: 21-32.

Desai, S. D., Chugh, D., & Brief, A. P. 2014. The implications of marriage structure for men's workplace

attitudes, beliefs, and behaviors toward women. *Administrative Science Quarterly*, 59: 330-365.

Detert, J. R., Trevino, L. K., & Sweitzer, V. L. 2008. Moral disengagement in ethical decision making: A

study of antecedents and outcomes. *Journal of Applied Psychology*, 93: 374-391.

Dipboye, R. L., & Halverson, S. K. 2004. Subtle (and not so subtle) discrimination in organizations. In R.

- W. Griffin & A. M. O’Leary-Kelly (Eds.), *The dark side of organizational behavior*. San Francisco: John Wiley & Sons.
- DiPrete, T. A., & Soule W. T. 1988. Gender and promotion in segmented job ladder systems. *American Sociological Review*, 53: 26-40.
- Dreber, A., & Johannesson, M. 2008. Gender differences in deception. *Economics Letters*, 99: 197–199.
- Dreher, G. F., & Cox, T. H. 2000. Labor market mobility and cash compensation: The moderating effects of race and gender. *Academy of Management Journal*, 43: 890-900.
- Duguid, M. 2011. Female tokens in high-prestige work groups: Catalysts or inhibitors? *Organizational Behavior and Human Decision Processes*, 116: 104-115.
- Duguid, M. M., Lloyd, D. L., & Tolbert, P. S. 2012. The impact of categorical status, numeric representation, and work group prestige on preference for demographically similar others: A value threat approach. *Organization Science*, 23: 386-401.
- Durkheim, E. 1925/1976. *The elementary forms of the religious life* (2nd ed.). London: Allen and Unwin.
- Eagly, A. H., Diekmann, A. B., Johannesen-Schmidt, M. C., & Koenig, A. M. 2004. Gender gaps in sociopolitical attitudes: A social psychological analysis. *Journal of Personality and Social Psychology*, 87: 796-816.
- Eagly, A. H., & Karau, S. J. 2002. Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109: 573-598.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. 1992. Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111: 3-22.
- 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., Wood, W., & Diekmann, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes, & H. M. Trautner (Eds.), *The developmental social psychology of gender* (pp. 123-174). Mahwah, NJ: Lawrence Erlbaum.

- Ely, R. J. 1995. The power in demography: Women's social constructions of gender identity at work. *Academy of Management Journal*, 38, 589-634.
- Ely, R. J., & Meyerson, D. E. 2000. Theories of gender in organizations: A new approach to organizational analysis and change. *Research in Organizational Behavior*, 22, 103-151.
- Ely, R., & Padavic, I. 2007. A feminist analysis of organizational research on sex differences. *Academy of Management Review*, 32: 1121-1143.
- Ely, R., Stone, P., & Ammerman, C. 2014. Rethink what you 'know' about high-achieving women. *Harvard Business Review*, 92: 101-109.
- Feldman, M. S., & March, J. G. 1981. Information in organizations as signal and symbol. *Administrative Science Quarterly*, 26: 171-186.
- Fernandez-Mateo, I. 2009. Cumulative gender disadvantage in contract employment. *American Journal of Sociology*, 114: 871-923.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. 2002. A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82: 878-902.
- Flynn, F. J., & Wiltermuth, S. S. 2010. Who's with me? False consensus, brokerage, and ethical decision making in organizations. *Academy of Management Journal*, 53: 1074-1089.
- Fragale, A. R., Rosen, B., Xu, C., & Merideth, I. 2009. The higher they are, the harder they fall: The effects of wrongdoer status on observer punishment recommendations and intentionality attributions. *Organizational Behavior and Human Decision Processes*, 108: 53-65.
- Franke, G. R., Crown, D. F., & Spake, D. F. 1997. Gender differences in ethical perceptions of business practices: A social role theory perspective. *Journal of Applied Psychology*, 82: 920-934.
- Gilligan, C. 1982. *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.
- Glick, P., & Fiske, S. T. 1996. The ambivalent sexism inventory: Differentiating hostile and benevolent

- sexism. *Journal of Personality and Social Psychology*, 70: 491–512.
- Goldin, C. & Rouse, C. 2000. Orchestrating impartiality: The impact of ‘blind’ auditions on female musicians.” *American Economic Review*, 90: 715-41.
- 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.
- Goodwin, G. P. 2015. Moral character in person perception. *Current Directions in Psychological Science*, 24: 38-44.
- Gorman, E. H. 2005. Gender stereotypes, same-gender preferences, and organizational variation in the hiring of women: Evidence from law firms. *American Sociological Review*, 70: 702-28.
- Haidt, J. 2007. The new synthesis in moral psychology. *Science*, 316: 998-1002.
- 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.
- Joshi, A. 2014. By whom and when is women’s expertise recognized? The interactive effects of gender and education in science and engineering teams. *Administrative Science Quarterly*, 59: 202–239.
- Kennedy, J. A., & Kray, L. J. 2013. Who is willing to sacrifice ethical values for money and social status?: Gender differences in reactions to ethical compromises. *Social Psychological and Personality Science*, 5: 52–59.
- Kim, P. H., Ferrin, D. L., Cooper, C. D., and Dirks, K. T. 2004. Removing the shadow of suspicion: The effects of apology vs. denial for repairing competence-versus integrity-based trust violations. *Journal of Applied Psychology*, 89: 104-118.
- Konrad, A. M., & Pfeffer, J. 1991. Understanding the hiring of women and minorities in educational institutions. *Sociology of Education*, 64: 141-157.
- Kray, L. J., & Haselhuhn, M. P. 2012. Male pragmatism in negotiators’ ethical reasoning. *Journal of*

- Experimental Social Psychology*, 48: 1124–1131.
- Kray, L. J., Kennedy, J. A., & Van Zant, A. B. 2014. Not competent enough to know the difference? Gender stereotypes about women's ease of being misled predict negotiator deception. *Organizational Behavior and Human Decision Processes*, 125: 61–72.
- Kulis, S. 1997. Gender segregation among college and university employees. *Sociology of Education*, 70: 151-173.
- Lerner, J. S., & Tetlock, P. E. 1999. Accounting for the effects of accountability. *Psychological Bulletin*, 125: 255-275.
- McGinn, K. L., & Milkman, K. L. 2013. Looking up and looking out: Career mobility effects of demographic similarity among professionals. *Organization Science*, 24: 1041-1060.
- Milkman, K. L., Akinola, M., & Chugh, D. 2012. Temporal distance and discrimination: An audit study in academia. *Psychological Science*, 23: 710-717.
- Mooijman, M., van Dijk, W. W., Ellemers, N., & van Dijk, E. 2015. Why leaders punish: A power perspective. *Journal of Personality and Social Psychology*, 109: 75-89.
- Olson, C. A., & Becker, B. E. 1983. Sex discrimination in the promotion process. *Industrial and Labor Relations Review*, 36: 624-41.
- Pfeffer, J., & Davis-Blake, A. 1987. The effect of the proportion of women on salaries: The case of college administrators. *Administrative Science Quarterly*, 32: 1-24.
- Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, 26: 269–281.
- Rehg, M. T., Miceli, M. P., Near, J. P., & Van Scotter, J. R. 2008. Antecedents and outcomes of retaliation against whistleblowers: Gender differences and power relationships. *Organization Science*, 19: 221-240.
- Reskin, B. F. 2000. The proximate causes of employment discrimination. *Contemporary Sociology*, 29:

319-28.

- Ridgeway, C. L. 2001. Gender, status, and leadership. *Journal of Social Issues*, 57: 637-655.
- Robinson, R. J., Lewicki, R. J., & Donahue, E. M. 2000. Extending and testing a five factor model of ethical and unethical bargaining tactics: Introducing the SINS scale. *Journal of Organizational Behavior*, 21: 649-664.
- Rucker, D. D., Polifroni, M., Tetlock, P. E., & Scott, A. L. 2004. On the assignment of punishment: The impact of general-societal threat and the moderating role of severity. *Personality and Social Psychology Bulletin*, 30: 673-684.
- Rudman, L. A. 1998. Self-promotion as a risk factor for women: The costs benefits of counterstereotypical impression management. *Journal of Personality Social Psychology*, 74: 629-645.
- Rudman, L., & Fairchild, K. 2004. Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, 87: 157-176.
- Rudman, L. A., & Phelan, J. E. 2008. Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, 28: 61-79.
- Ryan, M. K., & Haslam, S. A. 2007. The glass cliff: Exploring the dynamics surrounding the appointment of women to precarious leadership positions. *Academy of Management Review*, 32: 549-572.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. 2011. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22: 1359-1366.
- Tetlock, P. E. 2002. Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. 2002. *Psychological Review*, 109: 451-471.
- Tetlock, P. E., Visser, P. S., Singh, R., Polifroni, M., Scott, A., Elson, S. B., Mazzocco, P., & Rescober, P. 2007. People as intuitive prosecutors: The impact of social-control goals on attributions of responsibility. *Journal of Experimental Social Psychology*, 43: 195-209.

- Tetlock, P. E., & Mitchell, G. 2009. Implicit bias and accountability systems: What must organizations do to prevent discrimination? *Research in Organizational Behavior*, 29: 3-38.
- Trevino, L. K. 1992. The social effects of punishment in organizations: A justice perspective. *Academy of Management Review*, 17: 647-676.
- Tyler, T. R. 1990. *Why people obey the law*. New Haven, CT: Yale University Press.
- Uhlmann, E. L., & Cohen, G. L. 2005. Constructed criteria: Redefining merit to justify discrimination. *Psychological Science*, 16: 474-480.
- Vadera, A. K., & Pratt, M. G. 2013. Love, hate, ambivalence, or indifference? A conceptual examination of workplace crimes and organizational identification. *Organization Science*, 24: 172-188.
- Waytz, A. Dungan, J., & Young, L. 2013. The whistleblower's dilemma and the fairness-loyalty tradeoff. *Journal of Experimental Social Psychology*, 49: 1027-1033.
- Williams, J. E., & Best, D. L. 1982. *Measuring sex stereotypes: A thirty nation study*. Newbury Park, CA: Sage.
- Williams, M. J., & Tiedens, L. Z. 2016. The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin*, 142, 165-197.
- Wiltermuth, S. S., & Flynn, F. J. 2013. Power, moral clarity, and punishment in the workplace. *Academy of Management Journal*, 56: 1002-1023.
- Wright, E. O., Baxter, J., & Birkelund, G. E. 1995. The gender gap in workplace authority: A cross-national study. *American Sociological Review*, 60: 407-435.

TABLE 1
Descriptive Statistics and Correlations among Variables in Study 3

Variable	N	M	SD	1	2	3	4	5
1. Female attorney gender	482	0.17	0.37	--				
2. Punishment severity	482	1.89	1.31	.08	--			
3. Total rules violated	482	3.13	2.13	.02	.32***	--		
4. Attorneys per capita	482	22.78	56.95	-.07	.18***	.01	--	
5. Number of judges on panel	270	5.36	2.02	-.01	-.37	-.07	-.37***	--
6. Percentage of female judges	270	0.25	0.19	-.11	.16*	.05	.17**	.30***

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

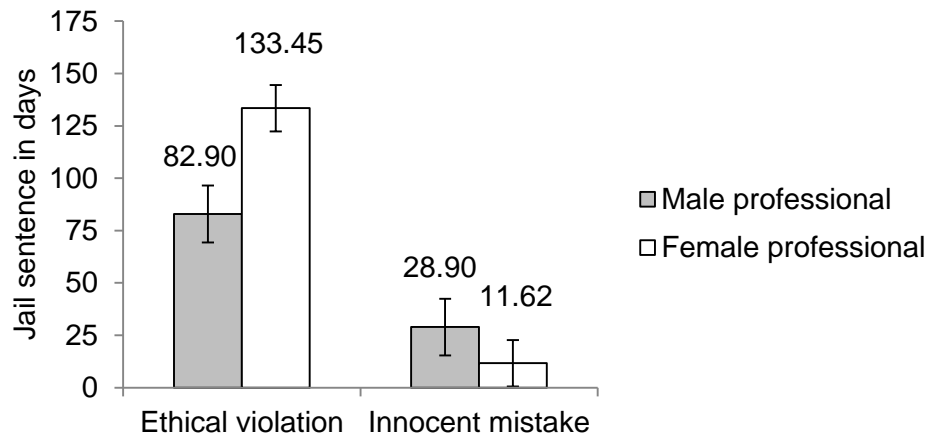
TABLE 2
Regression Models Predicting the Severity of Punishment in Study 3

	Ordinal Logistic Regressions of Categorical Punishment			Probit Regressions of Likelihood of Disbarment	
	Model 1	Model 2	Model 3	Model 4	Model 5
<i>Independent Variables</i>					
Female Gender	0.424*	1.046**	2.110***	0.567*	1.383***
	(0.18)	(0.37)	(0.58)	(0.26)	(0.33)
Female Gender x Percentage Female Judges			-4.867*		-4.066**
			(2.08)		(1.30)
<i>Control Variables</i>					
Total Rules Violated	0.408+	0.587+	0.576+	0.683***	0.673**
	(0.24)	(0.31)	(0.30)	(0.21)	(0.21)
Attorneys Per Capita	0.006***	0.008***	0.008***	0.004***	0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Number of Judges on Panel		0.204*	0.233*	0.131+	0.167*
		(0.09)	(0.09)	(0.07)	(0.07)
Percentage of Female Judges on Panel		0.229	0.628	0.006	0.375
		(0.99)	(0.91)	(0.78)	(0.76)
Fixed Effects for Rules Violated	YES	YES	YES	YES	YES
Fixed Effects for State	YES	YES	YES	YES	YES
Constant	^^	^^	^^	-2.937***	-3.260***
				(0.67)	(0.65)
Observations	482	270	270	270	270
Log Pseudolikelihood	-586.38	-303.19	-300.46	-119.97	-116.92
Pseudo R ²	0.1	0.14	0.15	0.22	0.24

Robust standard errors, clustered by state, in parentheses.

^^Because ordinal logistic regression results in multiple intercepts, we have omitted the constant for these models.

+ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

FIGURE 1**Punishment Severity in Study 2****FIGURE 2****Predicted Likelihood of Disbarment by Attorney's Gender in Study 3**