

# Hierarchy-Legitimizing Ideologies Reduce Behavioral Obligations and Blame for Implicit Attitudes and Resulting Discrimination

**Liz Redford - Kate A. Ratliff**

**Abstract** Three preregistered studies investigated people's judgments of whether someone with implicit racial bias is obligated to change their bias and to avoid discrimination based on that bias. Two studies showed that hierarchy-legitimizing ideologies—Belief in a Just World, Social Dominance Orientation, and political conservatism—predict lower obligation judgments. One study showed that hierarchy-legitimizing ideologies predicted greater protection of a potential discriminator; in another, they also predicted lower protection of a person who may be discriminated against. Lastly, one study showed that greater obligation judgments predicted greater blame of a person who discriminated based on implicit bias. Taken together, these four studies address how people's ideologies relate to their obligation judgments for implicit racial bias and how those obligation judgments are related to blame for discrimination resulting from implicit racial bias.

**Keywords** Morality Implicit Discrimination Ideology

## Introduction

People who intentionally discriminate incite blame and legal sanction (Bertelsen, 2012; Cushman, Young, & Hauser, 2006; Lagnado & Channon, 2008). But what about when people have automatic, unintended—implicit—racial biases? How much are those people to blame if their implicit bias results in discrimination? Implicit racial biases are related to harmful discrimination, even among those who consider themselves racially unbiased (O'Brien et al., 2010; Dasgupta, 2013). For example, implicit bias predicts voting behavior (Jost et al., 2009), costs people jobs (Jost et al., 2009) and medical help (Green et al., 2007), and it can threaten lives when police officers shoot Black men (Correll et al., 2007). Given implicit biases' potential to cause unintended harm, they raise interesting, important questions about responsibility and blame. What ought an implicitly biased person do? But the answers are more than just theoretically interesting. Interventions that attempt to change implicit bias or reduce their behavioral influence (e.g., Lai et al., 2014; Carnes et al., 2015 ; organizational diversity trainings) assume an obligation to change and to avoid discriminating. Understanding and changing whether their audiences agree is crucial to the success of these interventions.

The current research has four purposes. The first is to understand naive judgments about the obligations of a person with implicit racial bias: How much the person is obligated to change their implicit bias and how much the person is obligated to avoid discriminating based on their implicit bias. The second purpose is to test whether people's obligation judgments can be predicted by how much they endorse ideologies that legitimize social hierarchies. The present research also tests whether endorsement of

hierarchy-legitimizing ideologies predicts the importance people ascribe to the difficulty of changing implicit bias and avoiding its discriminatory effects, and whether they also predict importance ascribed to the potential harm to the person who may be discriminated against (potential discriminatee). Lastly, this research tests whether greater perceived obligation for implicit bias predicts greater blame when a target person discriminates.

## Judging Moral Obligations

In situations that pose a risk of harm, people perceive there to be a duty, or obligation, to lessen the potential of harm to others. For example, drivers are obligated to be vigilant to the road (Malle, Guglielmo, & Monroe, 2014). People agree that drivers are obligated to be vigilant, which is why laws impose duties on drivers—duties to exercise care in the form of reasonable speeds and maintaining control of the vehicle—all to avoid harm to others (*Levesque v. Anchor Motor Freight, Inc.*, 1987; *Thomason v. Willingham*, 1968). Some drivers, like those who operate public transit, owe even greater obligations: for example, to protect passengers from harm (*Restat 2d of Torts*, 1977). In the current research, the first purpose was to assess perceptions of the extent of obligations for an implicitly biased person. Specifically, in Study 1, we asked people to assess the extent of an individual's obligation to change their implicit bias and obligation to avoid discriminating based on their implicit bias.

People may judge obligations by considering several factors, factors that fall into one of two categories: either the potential harm of discrimination or the difficulty by which the obligations are met (as categorized in legal contexts for nondiscriminatory obligations; *U.S. v. Carroll Towing*, 1947). Factors related to potential harm include the *probability* and *severity* of the potential harm done by the implicitly biased person. Factors related to the difficulty of meeting obligations (to avoid discriminating or to change implicit bias) are the *burden* and the *practicality* of the obligation. For example, the difficulty of a driver's obligation to vigilance consists of how burdensome, and how practical, it is for the driver to be vigilant. A driver who fails to meet an obligation to be vigilant invites the risk of a car accident—a probable and severely harmful event. People may also consider whether meeting the obligation would lessen the potential harm (*reduced harm*) and whether the person foresees the potential harm (*foresight*).

In the current research, we test how important people perceive these variables burden, practicality, probability, severity, reduced harm, and foresight—to be for judging obligations for an implicitly biased person. A person who sees the obligation to avoid discriminating as minimally difficult, and sees the potential harm of discrimination as great, will impose a strict obligation to avoid discriminating. For example, a person who thinks it is easy to avoid discriminating based on implicit bias and who believes that implicit bias is likely to cause bias in police shootings (e.g., Correll et al., 2007) and life-threatening medical decisions (e.g., Green et al., 2007) may impose a greater obligation to avoid discriminating. A person who sees the obligation to avoid discriminating as very difficult, and sees the potential harm of discrimination to be low, will impose a more lenient obligation to avoid discriminating—if they impose any obligation at all. For example, a person who believes that implicit bias is impossible to control and who believes that implicit bias has only a small chance of causing minimal harm such as decreased eye contact (e.g., McConnell & Leibold, 2001) may impose only minimal, or

no obligation, to avoid discriminating. If so, then people who perceive difficulty to a potential discriminator to be more important than the potential harm target of discrimination (a “discriminatee”) should judge obligation to be lower.

No laws encode the moral obligations required of a person with implicit racial bias, so how might the average person judge the extent of these obligations? We expect that those who more strongly endorse group-based social hierarchies—the dominance of some group over others—will prioritize the difficulty to a potential discriminator over the potential harm to a potential discriminatee.

## **Obligation and Hierarchy-Legitimizing Ideologies**

We expect that those who more strongly endorse group-based social hierarchies will be more motivated to protect a potential discriminator from undue obligation than to protect a potential discriminatee from harm—in other words, to lessen potential harm. This is because potential discriminates are perceived as being relatively low in status—that is, low in the social hierarchy (e.g., Ridgeway & Diekema, 1992 ; Sekaquaptewa & Espinoza, 2004; Sidanius & Pratto, 1999 ). The status discrepancy in this situation makes it reasonable to expect hierarchy legitimation to predict reactions to it. Thus, the second purpose of this research is to test the hypothesis that those who more strongly endorse group-based social hierarchies will more strongly protect a potential discriminator and less strongly protect a potential discriminatee. To measure endorsement of group-based social hierarchies, we focus on three hierarchy-legitimizing ideologies: Social Dominance Orientation (SDO; Pratto, Sidanius, Stallworth, & Malle, 1994), Belief in a Just World (BJW; Lipkus, Dalbert, Siegler, 1996), and political conservatism. People high in SDO prefer unequal, hierarchical intergroup relations. Belief in a Just World reflects the motivation to believe that people generally get what they deserve and deserve what they get (e.g., Jost & Hunyady, 2005). It has often been conceptualized as a system-justifying (and thereby hierarchy-legitimizing) ideology (e.g., Jost & Hunyady, 2005; Kaiser, Dyrenforth, & Hagiwara, 2006; Kay & Jost, 2003; O’Brien & Major, 2005; Oldmeadow & Fiske, 2007), but when crossed with other variables like political efficacy, may not inevitably increase justification of inequality (Beierlein, Werner, Preiser, & Wermuth, 2011 ). Similarly, political conservatism legitimizes and enhances group hierarchies (Pratto et al., 1994), partly motivated by endorsement and rationalization of inequality (Jost et al., 2003). For example, conservatism is related to a preference for capitalism over socialism (e.g., Eysenck, 1971) and for hierarchical relationships (Wilson, 2013).

Other constructs like Right-Wing Authoritarianism (RWA) and ingroup identification are important and interesting predictors of prejudice and discrimination. However, the current research is focused on the *hierarchy legitimation* aspect of thoughts about implicit bias, an aspect much better captured by system-justifying constructs. For example, Social Dominance Orientation and constructs like RWA diverge significantly and predict prejudice for different reasons (Thomsen, Green, & Sidanius, 2008). RWA is uniquely related to prejudice against groups perceived to be dangerous (Duckitt & Sibley, 2007). However, people high in SDO are prejudiced against low-status groups because it legitimizes those groups’ low status (Duckitt & Sibley, 2007; Duckitt, 2009): exactly the sort of hierarchy legitimation relevant to the current research.

Although SDO, BJW, and political conservatism are unique constructs, they may

have similar implications for obligations regarding implicit bias. For these people, imposing obligations on a potential discriminator may seem to inappropriately challenge the legitimate and preferred racial hierarchy. They may care more about the difficulty and burden of meeting an obligation to avoid discriminating or change bias, and care less about the severity and harm to a potential discriminatee. Thus, imposing a harm-reducing obligation on a potential discriminator would not seem worth the difficulty and burden. If so, they would judge an implicitly biased potential discriminator to have lower obligations. Thus, a third purpose of this research is to test the hypothesis that those who more strongly endorse group-based social hierarchies will judge obligations to be lower. In Study 1 we measured participants' perceptions of the obligations of a person with implicit racial bias: How much the person is obligated to change their implicit bias and how much the person is obligated to avoid discriminating based on their implicit bias, and also tested the hypothesis that endorsing hierarchy-legitimizing ideologies will be related to lower obligation judgments.

In designing Study 1, we reasoned that people who endorse hierarchy legitimating ideologies would report lower obligation judgments because they would prioritize a potential discriminator over a potential discriminatee. Study 2 tested this reasoning: Whether people who legitimize hierarchies perceive difficulty to a potential discriminator to be more important than potential harm to a potential target of discrimination. Specifically, Study 2 tested the hypothesis that endorsement of hierarchy-legitimizing ideologies predicts greater motives to protect an implicitly biased potential discriminator and lower motives to protect a potential discriminatee. We measured how important participants thought the difficulty to a potential discriminator was (burden and practicality of obligation), and how important they thought the potential harm to a discriminatee was (severity and probability of discrimination), for judging obligations. We interpreted perceived importance of the difficulty to a potential discriminator (burden and practicality) as *potential -discriminator protection*. Thus, people who thought that the burden and practicality of meeting an obligation was very important for judging the extent of that obligation scored high on potential-discriminator protection. On the other hand, we interpreted perceived importance of the potential harm to a potential discriminatee (severity and probability) as *potential -discriminatee protection*.

## **Obligations and Blame**

Recall that drivers are obligated to be vigilant. Failure to be vigilant, especially when it causes a crash, invites blame. More generally, when someone fails to meet a harm-reducing obligation, thereby allowing the potential harm to happen, they are blamed (Brady, 1996 ; Redford & Ratliff, 2015 ). In other words, when the harm occurs, greater obligation leads to greater blame. The relationship between obligation and blame has mostly evaded empirical attention (Malle et al., 2014), but one exception shows that obligation judgments predict blame. In a series of studies, people read vignettes about a company manager who discriminates based on implicit racial bias. Those that reported that the manager had a greater obligation to foresee his discriminatory behavior judged him to be more blameworthy for discriminating (Redford & Ratliff, 2015). While these studies demonstrate a positive relationship between obligation and blame, they do so for only one form of obligation: the obligation to foresee discriminatory behavior. It may be

that different obligations—to change implicit bias or to avoid discriminating—also predict blame. If so, then people who judge an implicitly biased person to have greater obligations should attribute more blame to the implicitly biased person who discriminates. Thus, the fourth purpose of the current research was to test whether greater obligation to change implicit bias and greater obligation to avoid discriminating predict greater blaming of a person who discriminates based on implicit bias.

## Overview of the Present Research

Study 1 tested the hypothesis that endorsement of hierarchy-legitimizing ideologies (Social Dominance Orientation, Belief in a Just World, and political conservatism) predict lower judgments that a person with implicit bias is obligated to change their bias or is obligated to avoid discriminating. Participants read a vignette introducing implicit bias and obligations, and then answered questions about the extent of the person's obligations and about their own ideological preferences. Study 2 tested the hypothesis that hierarchy-legitimizing ideologies predict greater potential-discriminator protection and lower potential-discriminatee protection. Participants reported the importance of each of the six factors (severity, etc.) for each obligation judgment, and their ideological preferences. Study 3 combined and replicated Studies 1 and 2. Participants reported the extent of each obligation, the importance of each of the six factors for each obligation judgment, and their ideological preferences. Study 4 tested the hypothesis that greater perceptions that an implicitly biased person has certain behavioral obligations (to change their bias and prevent it from influencing their behavior) predict greater blame when the person discriminates due to their implicit bias. Taken together, these four studies address how people's ideologies relate to their obligation judgments for implicit bias and how those obligation judgments are related to blame.

## Study 1

### Method

#### *Participants*

We used G\*Power to conduct a power analysis with  $\alpha = 0.05$ , power = 0.95, and a small effect, for a regression analysis of three predictors on a continuous outcome. Attaining a small effect ( $R^2 = 0.04$ ) with those parameters requires 417 participants, which we chose as our target sample size. Because Project Implicit data collection does not stop immediately upon researcher request, the final sample size was larger. Participants were 570 White US citizen volunteers at the Project Implicit web site (<http://implicit.harvard.edu>; Nosek, Banaji, & Greenwald, 2002;  $M_{\text{age}} = 33.0$  years,  $SD = 14.2$ , 61.8 % women). The mean political orientation was 0.61 ( $SD = 1.74$ ) on a scale ranging from -3 = *Very conservative* to 3 = *Very liberal*. The current analyses focus on White Americans because they are likely to infer that potential discriminators, whose racial backgrounds are unspecified, are also White (Merritt & Harrison, 2006). This inference implies a racial status differential between potential discriminators and

potential discriminates, such that hierarchy-legitimizing concerns are relevant. In addition, given that European Americans show greater implicit bias against African Americans (Nosek et al., 2007), their judgments of obligations are more relevant to their own biases and behaviors.

### *Materials and Measures*

*Vignette* All participants read the same vignette, which introduced implicit bias and obligation and asked participants to prepare to make obligation judgments:

Many Americans consider themselves racially unbiased, but at the same time have negative automatic, gut-level feelings toward African Americans, which are relatively unintentional. In other words, it is possible that they associate African Americans with negativity without knowing it. Research shows that even if a person considers themselves racially unbiased, these negative automatic feelings could cause them to harmfully discriminate against African Americans.

When someone risks harming other people, they may have a duty or obligation to behave with a certain amount of care, caution, and attention. For example, when a person is driving a car, it is reasonable to expect them make an effort to pay attention to the road and look out for pedestrians and other drivers.

Now think about someone who considers themselves racially unbiased, but has negative automatic feelings toward African Americans that could cause them to discriminate. Next, you will answer questions about what duties or obligations you think this person has.

The content of the vignette was intentionally general in nature. Laypeople may hold vague and impoverished opinions about implicit attitudes, learned via popular press (e.g., that implicit attitudes = “unconscious racism” or “secret racism”; “How can we fix unconscious racism?” 2015; “Implicit Association Test, 2012). The same can hold for how people learn about the effect of implicit biases on discrimination. For example, the Frequently Asked Questions page at Project Implicit (<http://implicit.harvard.edu>) states that “... implicit biases can predict behavior. When we relax our active efforts to be egalitarian, our implicit biases can lead to discriminatory behavior, so it is critical to be mindful of this possibility if we want to avoid prejudice and discrimination.” Thus, the vignettes’ general nature was intended to address how laypeople really think and learn about implicit attitudes.

Materials and data from all studies and preregistrations of Studies 1–3 are available on the Open Science Framework at [https://osf.io/zj8an/?view\\_only=14a60e3fb95d4c009660fca33968262c](https://osf.io/zj8an/?view_only=14a60e3fb95d4c009660fca33968262c)

*Hierarchy-Legitimizing Ideologies* We measured Social Dominance Orientation using a Social Dominance Orientation scale ( $\alpha = 0.75$ ; Pratto et al., 1994), which contains six items such as *Inferior groups should stay in their place*. Participants responded on a 7-point scale from 1 = *Strongly disagree* to 7 = *Strongly agree*. We measured Belief in a

Just World using the Belief in a Just World scale, which includes eight items such as *I feel that people get what they deserve* ( $\alpha = 0.83$ ; Lipkus et al., 1996). Participants responded on a 7-point scale ranging from 1 = *Strongly disagree* to 7 = *Strongly agree*. We measured political orientation using a single item: *What is your political identity?* Participants responded on a 7-point scale coded as -3 = *Very conservative* to 3 = *Very liberal*. For analysis, political orientation was recoded into a conservatism variable, so that 1 = *Very liberal* and 7 = *Very conservative*.

*Obligation Judgments* Participants were first instructed to *Think about the person who considers themselves racially unbiased, but has negative automatic feelings toward African Americans that could cause them to discriminate*. We then measured obligation to avoid discriminating using a single item: *How obligated is this person to avoid discriminating based on their negative feelings?* and obligation to change negative feelings using a single item: *How obligated is this person to change their negative feelings?* Participants responded on a 5-point scale from 1 = *Not at all obligated* to 5 = *Extremely obligated*.

### *Procedure*

After random assignment to this study from the Project Implicit research pool, participants read the vignette introducing implicit bias. They then reported their perceptions of the target person's obligation to change their attitudes, and obligation to avoid discriminating, in randomized order. They then responded to measures of Social Dominance Orientation and Belief in a Just World, in randomized order, followed by the political orientation measure. Lastly, as specified in the preregistration, they completed an unrelated Implicit Association Test so that they could receive feedback, a requirement of the Project Implicit web site.

## **Results**

### *Obligation Judgments*

On average, participants thought that the target person was between *moderately* and *very* obligated to change their attitudes ( $M = 3.40$ ,  $SD = 1.24$ , where 3 = *Moderately obligated* and 4 = *Very obligated*), and *very* obligated to avoid discriminating ( $M = 3.92$ ,  $SD = 1.17$ ). Although we had no hypotheses regarding relative levels of the two types of obligation judgment, we chose to explore potential differences between them. A paired-samples  $t$  test showed that people judged there to be a significantly higher obligation to avoid discriminating than to change one's negative attitudes,  $t(555) = 11.32$ ,  $p < 0.001$ ,  $d = 0.96$ .

### *Main Analysis Strategy*

To test the hypotheses that hierarchy-legitimizing ideologies would predict both obligation judgments, we analyzed the data using two simultaneous regression analyses. Obligation to avoid discriminating and obligation to change implicit attitudes were each predicted by conservatism, Social Dominance Orientation, and Belief in a Just World. No

interactions were anticipated nor included in the model (as stated in the preregistration). Because the predictors were highly intercorrelated (see Table 2), we ran collinearity diagnostics. Multicollinearity was not a problem; tolerances ranged from 0.74 to 0.80 for the analyses reported, where 0.10 is recommended as the minimum level of tolerance (Tabachnick & Fidell, 2001).

### Main Analysis

*Obligation to Change Negative Attitudes* The overall model was significant,  $F(3,526) = 42.14$ ,  $p < 0.001$ ,  $R^2 = 0.19$ . Social Dominance Orientation (SDO) predicted obligation to change,  $b = -0.28$  (95 % CI - 0.18, - 0.38),  $\beta = -0.24$ ,  $SE = 0.05$ ,  $p < 0.001$ , such that greater SDO predicted lower obligation judgments (see Table 1). Belief in a Just World (BJW) predicted obligation to change,  $b = -0.21$  (95 % CI - 0.13, - 0.31),  $\beta = -0.19$ ,  $SE = 0.05$ ,  $p < 0.001$ , such that greater BJW predicted lower obligation judgments. Political conservatism predicted obligation to change,  $b = -0.10$  (95 % CI - 0.04, - 0.16),  $\beta = -0.14$ ,  $SE = 0.03$ ,  $p = 0.002$ , such that greater conservatism predicted lower obligation judgments.

**Table 1** Results from multiple regression analysis in Study 1

	$\beta$	$p$	SE
<i>Obligation to change negative attitudes</i>			
Social Dominance Orientation	-0.24	<0.001	0.05
Belief in a Just World	-0.19	<0.001	0.05
Political conservatism	-0.14	0.002	0.03
<i>Obligation to avoid discriminating</i>			
Social Dominance Orientation	-0.32	<0.001	0.05
Belief in a Just World	-0.10	0.030	0.05
Political conservatism	-0.13	0.005	0.03

$N = 570$ ,  $SE$  standard error,  $\beta$  standardized regression weights

*Obligation to Avoid Discriminating* The overall model was significant,  $F(3,529) = 41.97$ ,  $p < 0.001$ ,  $R^2 = 0.19$ . Social Dominance Orientation (SDO) predicted obligation to avoid discriminating,  $b = -0.35$  (95 % CI -0.25, -0.45),  $\beta = -0.32$ ,  $SE = 0.05$ ,  $p < 0.001$ , such that greater SDO predicted lower obligation judgments (see Table 1). Belief in a Just World (BJW) predicted obligation to avoid discriminating,  $b = -0.10$  (95 % CI -0.01, - 0.19),  $\beta = -0.10$ ,  $SE = 0.05$ ,  $p = 0.030$ , such that greater Belief in a Just World predicted lower obligation judgments. Political conservatism predicted obligation to avoid discriminating,  $b = -0.09$  (95 % CI -0.03, -0.15),  $\beta = -0.13$ ,  $SE = 0.03$ ,  $p = 0.005$ , such that greater conservatism predicted lower obligation judgments.

### Discussion

This study tested the hypothesis that endorsement of hierarchy-legitimizing ideologies would be negatively related to obligation judgments. As expected, Social Dominance Orientation, Belief in a Just World, and political conservatism predicted lower judgments of obligation to change negative implicit bias. In addition, Social Dominance Orientation

and political conservatism predicted judgments of obligation to avoid discriminating based on implicit bias. Hierarchy-legitimizing ideologies predicted a substantial part of obligation judgments: 19 % of the variance in obligation to avoid discriminating and 19 % of the variance in obligation to change bias.

In designing Study 1, we reasoned that hierarchy legitimation predicts lower obligation judgments because protecting a potential discriminatee seems less important than protecting a potential discriminator from undue obligation. In other words, for people who endorse hierarchy-legitimizing ideologies, imposing a harm-reducing obligation on a potential discriminator may not seem worth the difficulty and burden, which results in their judging the potential discriminator to have lower obligations. However, Study 1 measured only hierarchy legitimation and obligation judgments, not perceived importance of potential-discriminator and potential discriminatee protection. Study 2 bridged this gap, testing whether hierarchy legitimation predicts greater potential-discriminator protection and lower potential discriminatee protection.

## **Study 2**

Study 2 tested the hypothesis that hierarchy-legitimizing ideologies predict lower potential-discriminatee protection and greater potential-discriminator protection. To measure how motivated people are to protect a potential discriminatee, we asked them how important severity and likelihood of the potential harm caused by discrimination were for their obligation judgments. Thus, someone who thought severity and likelihood of the potential harm were very important would score high on potential-discriminatee protection. To measure how motivated people are to protect a potential discriminator, we asked them how important the burden and the practicality of meeting each obligation were for their obligation judgments.

## **Method**

### *Participants*

We used G\*Power to conduct a power analysis with  $\alpha = 0.05$ , power = 0.95, and a small effect, for a regression analysis of three predictors on a continuous outcome. As in Study 1, this power analysis indicated need for 417 participants, which we chose as our target sample size. Because Project Implicit studies do not stop immediately upon researcher request, the final sample size was greater than 417.

Participants were 502 White US citizen volunteers at the Project Implicit web site (M<sub>age</sub> = 34.4 years, SD = 14.9, 65.1 % women). The mean political orientation was 0.72 (SD = 1.70).

### *Materials and Measures*

*Vignette and Hierarchy-Legitimizing Ideologies* Participants read the same vignette as in Study 1. Also as in Study 1, we measured three hierarchy-legitimizing ideologies: Social Dominance Orientation ( $\alpha = 0.74$ ), Belief in a Just World ( $\alpha = 0.83$ ), and political orientation (single item). For analysis, political orientation was recoded into a

*conservatism* variable, so that 1 = *Very liberal* and 7 = *Very conservative*.

*Importance Ratings of Potential Obligation Determinants* We measured the perceived importance of six potential determinants of obligation, for both types of obligation (obligation to avoid discriminating and obligation to change bias), for a total of twelve items. Participants rated the importance of the following six potential obligation determinants: (1) Severity of the potential harm (*severity*), (2) Likelihood of the potential harm (*probability*), (3) The burden of meeting the obligation (*burden*), (4) The practicality of meeting the obligation (*practicality*), (5) Whether meeting the obligation would lessen the potential harm of discrimination (*reduced harm*), and (6) Whether the target person foresees the potential harm (*foresight*). For example, when rating the importance of the foreseeability of the potential harm for the obligation to avoid discriminating, the item read: *This person may or may not foresee the potential harm to African Americans. How important is whether or not the person foresees the potential harm for how much they are obligated to avoid discriminating based on their negative feelings?*

Participants responded on a five-point scale from 1 = *Not at all important* to 5 = *Extremely important*. Because the difficulty (burden and practicality) to potential discriminators include the burden and practicality of both types of obligations, they were combined: burden and practicality for both obligations (four items) were combined into a *potential-discriminator protection* variable ( $\alpha = 0.89$ ). Similarly, because potential harm (severity and probability) to potential discriminates encompass severity and probability of harm for both types of obligations, they were combined: severity and probability for both obligations (four items) were combined into a *potential-discriminatee protection* variable ( $\alpha = 0.87$ ). Importance of reduced harm for both obligations was combined into a general *reduced harm* variable, and importance of foresight for both obligations was combined into a general *foresight* variable.

### *Procedure*

As in Study 1, participants read the vignette about a person with implicit racial bias. They then reported the importance of six potential obligation determinants for that type of obligation. Thus, after reporting the obligation to change negative feelings, participants rated how important each of the six potential obligation determinants was for their judgment of the obligation to change negative feelings. Similarly, after reporting the obligation to avoid discriminating based on implicit attitudes, participants rated how important each of the six potential obligation determinants were for judging the obligation to avoid discriminating. Participants then responded to ideology measures and completed an unrelated Implicit Association Test.

## **Results**

### *Analysis Strategy*

To test the hypotheses that Social Dominance Orientation, Belief in a Just World, and political conservatism predict importance ratings of obligation determinants, we analyzed the data using two simultaneous regression analyses: one for potential discriminator

protection and one for potential-discriminatee protection. We were interested only in main effects, so the primary model did not include interactions (as stated in the preregistration). Also as specified in the preregistration, foresight and reduced harm were included as exploratory variables, for which we only report descriptive statistics. Because the predictors were highly intercorrelated (see Table 3), we ran collinearity diagnostics. Multicollinearity was not a problem; tolerances ranged from 0.75 to 0.89 for the analyses reported, where 0.10 is recommended as the minimum level of tolerance (Tabachnick & Fidell, 2001).

### *Potential-Discriminator Protection*

The overall model was significant,  $F(3,458) = 4.24, p = 0.006, R^2 = 0.03$ . Only one of the three predictor variables significantly predicted potential-discriminator protection: conservatism,  $b = 0.10$  (95 % CI 0.03, 0.16),  $\beta = 0.15, SE = 0.03, p = 0.003$ , such that greater conservatism predicted greater potential-discriminator protection (see Table 2). Contrary to expectations, Social Dominance Orientation did not predict potential-discriminator protection,  $b = 0.00$  (95 % CI -0.10, 0.10),  $\beta = 0.00, SE = 0.05, p = 0.96$ ; nor did Belief in a Just World,  $b = 0.03$  (95 % CI -0.07, 0.13),  $\beta = 0.03, SE = 0.03, p = 0.60$ .

### *Potential-Discriminatee Protection*

The overall model was not significant,  $F(3, 454) = 1.99, p = 0.12, R^2 = 0.01$ . Contrary to expectations, none of the three predictor variables significantly predicted potential discriminatee protection: Social Dominance Orientation did not predict potential-discriminatee protection,  $b = -0.05$  (95 % CI -0.17, 0.06),  $b = -0.05, SE = 0.06, p = 0.38$ ; nor did Belief in a Just World,  $b = 0.08$  (95 % CI -0.03, 0.20),  $\beta = 0.07, SE = 0.06, p = 0.16$ ; nor did conservatism,  $b = 0.05$  (95 % CI -0.02, 0.13),  $\beta = 0.08, SE = 0.04, p = 0.15$  (see Table 2).

**Table 2** Results from multiple regression analysis in Study 2

	$\beta$	$p$	SE
<b>Potential-discriminator protection</b>			
Social Dominance Orientation	0.00	0.96	0.05
Belief in a Just World	0.03	0.60	0.03
Political conservatism	0.15	0.003	0.03
<b>Potential-discriminatee protection</b>			
Social Dominance Orientation	-0.05	0.38	0.06
Belief in a Just World	0.07	0.16	0.06
Political conservatism	0.08	0.15	0.04

*N* = 502, *SE* standard error,  $\beta$  standardized regression weights

### *Importance of Foresight and Reduced Harm*

As specified in the preregistration, foresight and reduced harm were included as exploratory variables, for which we only report descriptive statistics. On average,

participants perceived an implicitly biased person's foresight of potential harm of discrimination (*foresight*) to be approximately moderately important ( $M = 3.22$ ,  $SD = 1.23$ , where 3 = *Moderately important*) for judging that person's obligations. They perceived whether meeting the obligation would lessen the potential harm of discrimination (*reduced harm*) to be between moderately and very important ( $M = 3.62$ ,  $SD = 1.15$ , where 4 = *Very important*). Although we had no hypotheses regarding relative levels of foresight and reduced harm, we chose to explore potential differences between them. A paired-samples  $t$  test showed that people thought that reduced harm was significantly more important for obligation judgments than was foresight,  $t(484) = 7.85$ ,  $p < 0.001$ ,  $d = 0.71$ .

## Discussion

Study 2 showed that as a group, hierarchy-legitimizing ideologies predict greater motives to protect an implicitly biased potential discriminator (burden and practicality of obligation). However, in terms of the three individual predictors, only political orientation significantly predicted potential-discriminator protection. Study 2 also showed that hierarchy-legitimizing ideologies did not predict lower motives to protect the person who may be discriminated against (severity and probability of discrimination).

While we stated in the preregistration that we expected foresight and reduced harm to be important, we did not specify a numeric value that would indicate importance. However, we interpret the average rating (between *Moderately important* and *Very important*) as an exploratory indication that foresight and reduced harm are important for obligation judgments. In addition, people thought that reduced harm was significantly more important for obligation judgments than was foresight. This supports our reasoning that obligations are designed to lessen risk of harm, because participants recognized that an obligation's harm-reducing power should weigh heavily in their obligation judgments.

## Study 3

Study 3 combined and replicated Studies 1 and 2. In Study 1, greater endorsement of hierarchy-legitimizing ideologies (political orientation, Social Dominance Orientation, and Belief in a Just World) predicted greater perceptions that an implicitly biased person is obligated to avoid discriminating and obligated to change their implicit attitudes. Thus, we predicted that Study 3 would also show hierarchy-legitimizing ideologies (political orientation, Social Dominance Orientation, and Belief in a Just World) to predict both types of obligation judgment. Study 3 also tests the hypothesis, which was only partially supported in Study 2, that hierarchy-legitimizing ideology predicts greater potential-discriminator protection and lower potential-discriminatee protection.

## Method

### *Participants*

As specified in the preregistration, we conducted a power analysis using the effect size for the smallest significant model in Studies 1 and 2 ( $R^2 = 0.03$  in Study 2). We used

G\*Power to conduct a power analysis with  $\alpha = 0.05$  and power = 0.95, for a regression analysis of three predictors on a continuous outcome. This analysis indicated need for 560 participants. Because Project Implicit data collection does not stop immediately upon researcher request, the final sample size was larger. Participants were 618 White US citizen volunteers at the Project Implicit web site (Mage = 34.7 years, SD = 15.7, 63.1 % women). The mean political orientation was 1.09 (SD = 1.71).

### *Materials and Measures*

*Vignette* All participants read the same vignette as in Studies 1 and 2, which introduced implicit bias and obligation.

*Hierarchy-Legitimizing Ideologies* As in Studies 1 and 2, we measured Social Dominance Orientation ( $\alpha = 0.73$ ), Belief in a Just World ( $\alpha = 0.85$ ), and political orientation. For analysis, political orientation was recoded into a *conservatism* variable, so that 1 = *Very liberal* and 7 = *Very conservative*.

*Obligation Judgments* As in Study 1, we measured obligation to avoid discriminating and obligation to change negative feelings using one item each.

*Importance Ratings of Potential Obligation Determinants* As in Study 2, we measured importance ratings for six potential obligation determinants for each of the two types of obligation, for a total of twelve items. Participants rated the importance of the same potential obligation determinants as in Study 2: severity, probability, burden, practicality, foresight, and reduced harm.

### *Procedure*

Participants read the vignette introducing implicit bias and moral obligation. They then reported the extent of that person's obligations and the importance of six potential obligation determinants for that type of obligation, followed by measures of ideology and an unrelated Implicit Association Test.

## **Results**

### *Obligation Judgments*

On average, participants thought that the target person was between *moderately* and *very* obligated to change their attitudes ( $M = 3.75$ ,  $SD = 1.14$ , where 3 = *Moderately obligated* and 4 = *Very obligated*), and approximately very obligated to avoid discriminating ( $M = 4.11$ ,  $SD = 1.02$ ). Although we had no hypotheses regarding relative levels of the two types of obligation judgment, we chose to explore potential differences between them. A paired-samples t test showed that people thought there was a significantly higher obligation to avoid discriminating than to change negative attitudes,  $t(606) = 9.11$ ,  $p < 0.001$ ,  $d = 0.74$ .

### Predicting Obligation Judgments

*Analysis Strategy* The primary hypotheses were tested with two simultaneous regression analyses in which each type of obligation was predicted by conservatism, Social Dominance Orientation, and Belief in a Just World. We were interested only in main effects, so the primary model did not include interactions. Because the predictors were highly intercorrelated (see Table 4), we ran collinearity diagnostics. Multicollinearity was not a problem; tolerances ranged from 0.70 to 0.81 for the analyses reported, where 0.10 is recommended as the minimum level of tolerance (Tabachnick and Fidell 2001).

**Table 3** Results from multiple regression analysis in Study 3

	$\beta$	$p$	SE
<b>Obligation to change negative attitudes</b>			
Social Dominance Orientation	-0.32	<0.001	0.05
Belief in a Just World	-0.14	0.001	0.04
Political conservatism	-0.14	0.001	0.03
<b>Obligation to avoid discriminating</b>			
Social Dominance Orientation	-0.37	< .001	0.04
Belief in a Just World	-0.22	0.030	0.04
Political conservatism	-0.04	0.38	0.03
<b>Potential-discriminator protection</b>			
Social Dominance Orientation	0.13	0.004	0.05
Belief in a Just World	0.08	0.10	0.05
Political conservatism	0.17	0.001	0.03
<b>Potential-discriminatee protection</b>			
Social Dominance Orientation	0.03	0.57	0.06
Belief in a Just World	0.13	0.005	0.06
Political conservatism	0.06	0.21	0.04

$N = 618$ , SE standard error,  $\beta$  standardized regression weights

*Obligation to Change Negative Attitudes* The overall model was significant,  $F(3,582) = 56.70$ ,  $p < 0.001$ ,  $R^2 = 0.23$ . As expected, all three predictor variables significantly predicted obligation to change negative attitudes (see Table 3).

Social Dominance Orientation predicted obligation to change negative attitudes,  $b = -0.37$  (95 % CI -0.28, -0.46),  $\beta = -0.32$ , SE = 0.05,  $p < 0.001$ , such that greater Social Dominance Orientation predicted lower obligation judgments. Belief in a Just World predicted obligation to change negative attitudes,  $b = -0.14$  (95 % CI -0.06, -0.23),  $\beta = -0.14$ , SE = 0.04,  $p = 0.001$ , such that greater Belief in a Just World predicted lower obligation judgments. Conservatism predicted obligation to change negative attitudes,  $b = -0.09$  (95 % CI -0.04, -0.15),  $\beta = -0.14$ , SE = 0.03,  $p = 0.001$ , such that greater conservatism predicted lower obligation judgments.

*Obligation to Avoid Discriminating* The overall model was significant,  $F(3,579) = 66.93$ ,  $p < 0.001$ ,  $R^2 = 0.26$ . Two of the three predictor variables significantly predicted obligation to avoid discriminating (see Table 3). Social Dominance Orientation predicted obligation to avoid discriminating,  $b = -0.38$  (95 % CI -0.30, -0.46),  $\beta = -0.37$ , SE = 0.04,  $p < 0.001$ , such that greater Social Dominance Orientation predicted lower obligation

judgments. Belief in a Just World predicted obligation to avoid discriminating,  $b = -0.20$  (95 % CI -0.12, -0.27),  $\beta = -0.22$ , SE = 0.04,  $p = 0.03$ , such that greater Belief in a Just World predicted lower obligation judgments. Unexpectedly, conservatism did not predict obligation to avoid discriminating,  $b = -0.02$  (95 % CI -0.07, 0.03),  $\beta = -0.04$ , SE = 0.03,  $p = 0.38$ .

### *Predicting Potential-Discriminator and Potential-discriminatee Protection*

*Analysis Strategy* As in Study 2, to test the hypotheses that Social Dominance Orientation, Belief in a Just World, and political conservatism predict greater potential-discriminator protection and lower potential-discriminatee protection, we analyzed the data using two simultaneous regression analyses: one for potential discriminator protection and one for potential-discriminatee protection. We were interested only in main effects, so the primary model did not include interactions (as stated in the preregistration). Also as specified in the preregistration, foresight and reduced harm were included as exploratory variables, for which we only report descriptive statistics.

*Outcome Computation* As in Study 2, burden and practicality for both obligations had high interitem correlation,  $\alpha = 0.89$ , and were combined into a *potential discriminator protection* variable. Also as in Study 2, severity and likelihood for both obligations had high interitem correlation,  $\alpha = 0.87$ , and were combined into a *potential-discriminatee protection* variable. Importance of reduced harm for both obligations was combined, and importance of foresight for both obligations was combined. Importance of *reduced harm* for both obligations was combined into a general reduced harm variable, and importance of foresight for both obligations was combined into a general *foresight* variable.

*Potential-Discriminator Protection* The overall model was significant,  $F(3,580) = 17.62$ ,  $p < 0.001$ ,  $R^2 = 0.08$ . Two of the three predictor variables significantly predicted potential-discriminator protection: conservatism,  $b = 0.11$  (95 % CI 0.05, 0.18),  $\beta = 0.17$ , SE = 0.03,  $p = 0.001$ , such that greater conservatism predicted greater potential-discriminator protection, and Social Dominance Orientation,  $b = 0.15$  (95 % CI 0.05, 0.25),  $\beta = 0.13$ , SE = 0.05,  $p = 0.004$ , such that greater Social Dominance Orientation predicted greater potential-discriminator protection (see Table 3). Belief in a Just World did not predict potential-discriminator protection,  $b = 0.08$  (95 % CI - 0.01, 0.25),  $\beta = 0.08$ , SE = 0.05,  $p = 0.10$ .

*Potential-Discriminatee Protection* The overall model was significant,  $F(3,580) = 6.40$ ,  $p < 0.001$ ,  $R^2 = 0.03$ . Only one of the three predictor variables significantly predicted potential-discriminatee protection: Belief in a Just World,  $b = 0.15$  (95 % CI 0.05, 0.26),  $\beta = 0.13$ , SE = 0.06,  $p = 0.005$ , such that greater Belief in a Just World predicted greater potential-discriminatee protection (see Table 3). Social Dominance Orientation did not predict potential-discriminatee protection,  $b = 0.04$  (95 % CI - 0.09, 0.15),  $\beta = 0.03$ , SE = 0.06,  $p = 0.57$ ; nor did conservatism,  $b = 0.05$  (95 % CI - 0.03, 0.12),  $\beta = 0.06$ , SE = 0.04,  $p = 0.21$ .

*Importance of Foresight and Reduced Harm* As specified in the preregistration, foresight

and reduced harm were included as exploratory variables, for which we only report descriptive statistics. On average, participants perceived an implicitly biased person's foresight of potential harm of discrimination (*foresight*) to be moderately important ( $M = 2.94$ ,  $SD = 1.32$ , where 3 = *Moderately important*) for judging that person's obligations. They perceived whether meeting the obligation would lessen the potential harm of discrimination (*reduced harm*) to be between moderately and very important ( $M = 3.53$ ,  $SD = 1.26$ , where 4 = *Very important*). A paired-samples  $t$  test showed that people thought that reduced harm was significantly more important for obligation judgments than was foresight,  $t(600) = 11.10$ ,  $p < 0.001$ ,  $d = 0.91$ .

## Discussion

The present study combined and replicated Studies 1 and 2. Replicating the results of Study 1, it showed that hierarchy-legitimizing ideologies (political conservatism, Social Dominance Orientation, and Belief in a Just World) predict perceptions that an implicitly biased person is obligated to avoid discriminating and obligated to change their implicit attitudes. Social Dominance Orientation, Belief in a Just World, and political conservatism predicted lower obligation to change negative attitudes, and Social Dominance Orientation and Belief in a Just World predicted obligation to avoid discriminating. The variance of obligation judgments explained by hierarchy-legitimizing ideology is similar to that of Study 1: Hierarchy legitimation predicted 26 % of the variance in obligation to avoid discriminating and 23 % of the variance in obligation to change bias. These large effect sizes suggest that hierarchy-legitimizing ideologies are an important factor in judging moral obligation for discrimination based on implicit bias. The results of this study, along with Study 1, support our reasoning that people who endorse social hierarchies are more lenient in judging obligations of an implicitly biased potential discriminator.

The current study also tested whether hierarchy-legitimizing ideology predicts greater discriminator protection and lower discriminatee protection. However, the results did not exactly replicate the results of the previous study. In Study 2, political conservatism predicted greater potential-discriminator protection, but the model was not significant for potential-discriminatee protection. That is, SDO, BJW, and conservatism did not predict perceived importance of severity and probability of discrimination for obligation judgments. In this study, however, political orientation *and* Social Dominance Orientation predicted greater potential discriminator protection. In addition, the model was significant for potential discriminatee protection; Belief in a Just World predicted greater potential discriminatee protection. This last relationship was opposite of that expected. Effect sizes were small:  $R^2 = 0.03$  in Study 2, and 0.08 and 0.03 in Study 3. Thus, while most of the observed relationships were in the predicted direction, they were small and unstable.

While we stated in the preregistration that we expected foresight and reduced harm to be important, we did not specify a numeric value that would indicate importance. However, as in Study 2, we interpret the average rating (between *Moderately important* and *Very important*) as an exploratory indication that foresight and reduced harm are important for obligation judgments. Again, the greater importance of reduced harm than foresight supports our reasoning that obligations are designed to lessen risk of harm.

## Study 4

Studies 1, 2, and 3 addressed how people's ideologies predict their judgments of an implicitly biased person's obligations to change their bias and avoid discriminating. Legally, obligations lead to blame if a risked harm actually happens (Brady, 1996). Psychologically, as perceived obligation increases, perceived blame also increases (Redford & Ratliff, 2015). Thus, when an implicitly biased person actually discriminates, people who judge the person to have greater obligations to change and to avoid discrimination should attribute more blame to the person.

### Method

#### *Participants*

A G\*Power power analysis with  $\alpha = 0.05$ , power = 0.95, and a small effect, for a regression analysis of three predictors on a continuous outcome, indicated a need for 417 participants, which we chose as our target sample size. We miscalculated when we should stop data collection and ended up with a slightly smaller sample size. Participants were 410 White US citizen volunteers at the Project Implicit web site (Mage = 34.2 years, SD = 14.4, 70.5 % women). The mean political orientation was 0.73 (SD = 1.63).

#### *Materials and Measures*

*Vignette* All participants read the same vignette as in Studies 1, 2, and 3, which introduced implicit bias and obligation.

*Obligation Judgments* As in Studies 1 and 2, we measured obligation to avoid discriminating and obligation to change negative feelings using one item each.

*Blame* Participants were asked to imagine that as a *result of their negative automatic feelings, the person [with implicit bias] discriminates against African Americans*. To measure how much they blamed the target person, we used an adapted form of an existing moral criticism questionnaire (Redford & Ratliff, 2015). The measure was adapted from a four-item measure (Cameron, Payne, & Knobe, 2010) which assessed the extent to which participants saw a person as "morally responsible," "should be punished," "should [not] be blamed," and "should [not] be held accountable" for discriminating based on implicit attitudes. This four-item measure was adapted to eight for Redford and Ratliff (2015), in which it predicted awareness- and obligation-based blame for discrimination based on implicit attitudes. In the current research, we retain the four original items as well as: how much the person "should be considered immoral," "should be judged," and "is at fault," for a total of seven items ( $\alpha = 0.84$ ). These items improve the internal consistency of the measure (previous  $\alpha = 0.65$ ) and to address the agent-evaluative aspect of blame (Malle et al., 2012), by which perceivers judge the moral character of an actor (Uhlmann, Pizarro, & Diermeier, 2015), and which may be especially relevant in judgments of racial bias (Uhlmann, Zhu, & Diermeier, 2014).

## Procedure

Participants read the vignette and reported their obligation judgments. They then reported how much they blamed the target person for discriminating, and lastly completed an unrelated Implicit Association Test.

## Results

### Obligation Judgments

On average, participants thought that the target person was between moderately and very obligated to change their attitudes ( $M = 3.54$ ,  $SD = 1.20$ , where 3 = *Moderately obligated* and 4 = *Very obligated*), and very obligated to avoid discriminating ( $M = 4.06$ ,  $SD = 1.09$ ; see Table 1 for means and standard deviations of measures in all studies). People judged there to be a significantly higher obligation to avoid discriminating than to change negative attitudes,  $t(402) = 8.87$ ,  $p < 0.001$ ,  $d = 0.89$ .

### Predicting Blame

*Analysis Strategy* To test the prediction that obligation judgments would predict blame, we conducted a simultaneous multiple regression entering each obligation judgment, and their interaction, as predictors. We expected only main effects of the predictor variables, but planned to explore interactions between the predictors. Because the predictors were highly intercorrelated (see Table 4), we ran collinearity diagnostics. Multicollinearity was not a problem; tolerances ranged from 0.67 to 0.83 for the analyses reported, where 0.10 is recommended as the minimum level of tolerance (Tabachnick & Fidell, 2001).

*Main Analysis* The model was significant,  $F(3, 391) = 11.84$ ,  $p < 0.001$ ,  $R^2 = 0.08$ . Obligation to avoid discriminating significantly predicted blame,  $b = 0.16$  (95 % CI 0.08, 0.23),  $\beta = 0.24$ ,  $SE = 0.04$ ,  $p < 0.001$  (see Table 4). Obligation to change bias also significantly predicted blame,  $b = 0.06$  (95 % CI 0.002, 0.13),  $\beta = 0.11$ ,  $SE = 0.03$ ,  $p = 0.042$ . The interaction of obligation to avoid and obligation to change did not significantly predict blame,  $b = 0.04$  (95 % CI -0.002, 0.09),  $\beta = 0.10$ ,  $SE = 0.02$ ,  $p = 0.06$ .

**Table 4** Results from multiple regression analysis in Study 4

	$\beta$	$p$	SE
Obligation to change negative attitudes	0.11	0.042	0.03
Obligation to avoid discriminating	0.24	<0.001	0.04
Interaction of obligation to change with obligation to avoid	0.10	0.06	0.02

$N = 410$ .  $SE$  standard error

Dependent variable = blame for discrimination

## **Conclusion**

As expected, both greater obligation to change bias and greater obligation to avoid discriminating predicted more blame ascribed to an implicitly biased person who discriminates against African Americans.

## **General Discussion**

The current research examined people's perceptions of how much a person with implicit racial bias is obligated to change their bias and to avoid discriminating. As expected, people with greater hierarchy-legitimizing ideologies had lower perceptions of one's obligation to change negative attitudes and lower perceptions of one's obligation to avoid discriminating. That is, people who more strongly prefer the dominance of some social groups over others are more lenient in judging the obligations of an implicitly biased potential discriminator. We also tested the hypothesis that greater hierarchy-legitimizing ideologies would predict greater potential-discriminator protection and lower potential-discriminatee protection. Here, the findings were mixed; the observed relationships were small and unstable across the two studies that tested this hypothesis. Lastly, we tested the hypothesis that greater obligation to change implicit bias and to avoid discriminating would predict more blame of a person who discriminates based on implicit bias. As expected, people blamed a discriminator more when they believed more strongly in the discriminator's obligation to avoid discriminating or to change implicit bias.

Hierarchy-legitimizing ideologies consistently predicted large decreases in obligation judgments. This finding supports our argument that for people who endorse hierarchy-legitimizing ideologies, the racial hierarchy seems legitimate; thus, it seems inappropriate to challenge that hierarchy by imposing obligations on a high-status potential discriminator. More broadly, this ideology–obligation relationship suggests that moral reasoning can reflect preferences for hierarchical social structures. Thus, a broader implication of our findings is that for moral transgressions that support or challenge social hierarchies, moral evaluations reflect beliefs about whether the hierarchy is worth supporting (legitimate) or worth challenging (illegitimate). Moral reasoning about duties and obligations may be motivated by protecting transgressors when the transgression reinforces the preferred social hierarchy. However, moral and legal psychology, and the real world legal system, has not accounted for the fact that people judge obligations based on their personal preferences for hierarchies.

The finding that preferences for hierarchy predict obligation judgments can allow the public, policy-makers, and ethicists to decide whether the relationship of ideology to obligation judgments is fair or biased. For example, some may consider such ideologies grounds for juror exclusion in racial discrimination cases, while others may consider them reasonable decision inputs. The relationship between hierarchy preferences and obligation judgments can also inform interventions. Interventions that attempt to change implicit bias or reduce their behavioral influence (e.g., Lai et al., 2014 ; Carnes et al., 2015 ; organizational diversity trainings) assume an obligation to change and to avoid discriminating. The current results show that participant cooperation, and perhaps thereby intervention success, may depend on whether they endorse hierarchies. Future studies could test whether inducing egalitarian norms and preferences improves receptiveness to

such interventions.

There were some relationships of hierarchy-legitimizing ideologies to potential discriminator protection and potential-discriminatee protection, suggesting that at least sometimes, ideologies predict how a person judges obligations. We expected that the three hierarchy-legitimizing ideologies would uniformly predict greater potential-discriminator protection and lower potential-discriminatee protection. However, the results suggested that some ideologies might asymmetrically predict discriminator protection and discriminatee protection. For example, greater political conservatism significantly predicted greater potential-discriminator protection, but not lower potential-discriminatee protection. Thus, some aspects of hierarchy legitimation may exert hierarchy-maintaining effects via privileging high-status potential discriminators, while others may maintain hierarchies via devaluing low status potential discriminatees. For example, the current results suggest that for people who are more politically conservative, moral reasoning is at least partly aimed at protecting potential transgressors, but irrelevant to protecting potential targets of the transgression.

The last study supported our hypothesis that greater obligation would predict greater blame. People blamed a discriminator more when they believed that the discriminator's had a greater obligation to avoid discriminating or to change implicit bias. This finding supports our argument that people blame those who violate obligations, because failing to meet an obligation allows the risked harm to actually happen. The discriminator failed to meet the obligation to change bias (i.e., remains biased) and also failed to meet the obligation to avoid discriminating (i.e., discriminates).

This finding is one of the first to directly show the positive relationship between obligation and blame. This relationship is consistent with the Path Model of Blame, which theorizes obligation as a prerequisite for blame (Malle et al., 2014). It also suggests that people consider obligation in deciding culpability. For example, failure to meet a legal obligation is the basis of culpable negligence (Brady, 1996). Thus, in everyday moral reasoning, obligation judgments might predict gradations of blame for negligence, recklessness, and directly intended harm. For example, for a person who is highly obligated to avoid discriminating, their discriminating may be considered reckless or negligent. Alternatively, for a person who is not obligated to avoid discriminating, their discriminating may be considered merely accidental. If so, then obligation deserves a larger role in moral and legal psychology than it has so far been granted.

In addition, obligation to avoid discriminating more powerfully predicted blame than did obligation to change bias. This difference is meaningful. Given that the purpose of obligations is to mitigate risk of harm, it may be that the obligation more directly tied to preventing harm—avoiding discriminating—is more relevant to blame. This reasoning may also explain another of the current findings. Consistently, participants perceived obligation to avoid discriminating to be greater than obligation to change implicit attitudes ( $d_s = 0.96, 0.74, .89$ ). Such prioritization of the obligation to avoiding discriminating supports our reasoning that moral obligations are designed to prevent harm.

We reasoned that hierarchy-legitimizing ideologies would predict lower obligation judgments *because* they would motivate greater potential-discriminator protection and lower potential-discriminatee protection. However, we observed only small, unstable relationships between hierarchy legitimation and potential discriminator

protection and potential-discriminatee protection. This failure to fully support our hypotheses might be explained by the complexity of the topic and instructions to participants. We asked volunteers how important each determinant (e.g., the severity of harm) was for the obligation to avoid discriminating and also for the obligation to change bias. Volunteers may have struggled to respond accurately to the combination of the abstract concept of severity with other newly learned abstract concepts like implicit bias and obligation. Alternatively, it could be that hierarchy legitimation predicts obligation judgments, but not always via differential weighing of the obligation determinants (severity, etc.) that we included in this study. It may be that people consider different obligation determinants than those we tested or that they consider them in a way that does not relate to their consideration of difficulty and potential harm for potential discriminators and potential discriminatees.

Either way, these findings might be clarified by a clearer methodological approach, like concretely manipulating the six obligation determinants (severity, etc.). For example, severity of harm could be manipulated to be high or low: Researchers could describe discrimination resulting from implicit bias as either very harmful (i.e., death or injury) or less harmful (i.e., unfriendliness). The effect size of this severity manipulation on obligation judgments could then be interpreted as importance of severity. This method would bring its own limitations, among them being the reduction of continuous variables (i.e., severity) to dichotomous ones, thereby reducing power, accuracy, and generalizability.

Another potential explanation for the inconsistent findings is the use of Belief in a Just World (BJW) as a predictor. BJW is generally considered to be a system justifying (and thereby hierarchy-legitimizing) ideology (Jost & Hunyady, 2005). However, when crossed with other variables, it may not inevitably increase justification of inequality (Beierlein et al., 2011). This nuance may have contributed to BJW's inconsistent effects in our studies—effects which also tend to be weaker than those of Social Dominance Orientation, which more unambiguously represents hierarchy legitimation.

The current studies were based on racial discrimination, but their findings might apply to other moral violations that support or challenge hierarchies. In the current studies, the moral transgression—racial discrimination—supports the existing racial hierarchy. As expected, those who endorse existing hierarchies were more lenient in judging obligations of a potential discriminator. It could be that more generally, people who endorse hierarchy-legitimizing ideologies are lenient when the potential harm supports continued inequality. For example, in addition to judging low obligations for a person with implicit *racial* bias, they may also judge low obligations for a person with implicit preferences for men over women. In corollary, those who see hierarchies as illegitimate may impose obligations that protect people in the lower ranks. However, moral transgressions can also *challenge* the social order: Consider discrimination against Whites or Christians in the USA. In those cases, hierarchy legitimation may produce strictness rather than lenience. For example, someone who endorses the current racial hierarchy may prioritize protecting a potential White discriminatee at the expense of a potential Black discriminator.

In the current studies, we reasoned that people use instrumental, consequentialist, utilitarian reasoning to judge obligations. But people might also make *unconditional* obligation judgments that are free of considerations about net utility. Unconditional

obligation grounds deontology, a meta-ethical position whereby one chooses moral actions based on universal principles of moral duty. Unconditional moral judgments are also a major part of *moral mandates*, which are perceived to be self-justifying, independent of authority or cultural norms, and objective and universal (Skitka, 2010; 2014 ). If racial discrimination brings unconditional obligations, then the potential obligation determinants (e.g., severity) would not matter to participants: Their importance ratings would approach zero, limiting the explainable variance and potentially producing null results. Moreover, whether racial discrimination brings unconditional obligation judgments may depend on the judge's hierarchy legitimizing ideologies. For example, those low in hierarchy legitimation may be more likely to unconditionally obligate a potential discriminator to avoid discriminating. In corollary, a person high in hierarchy legitimation may be more likely to use utilitarian reasoning. Thus, future studies might consider when obligation judgments are conditional: for whom, for what moral violations, and in what contexts.

## Conclusions

Three studies show that ideologies predict obligation judgments, and one study shows that obligation judgments predict blame. These findings suggest that there may be an ideology–blame relationship mediated by obligation judgments. If so, then people who endorse hierarchy-legitimizing ideologies should blame a discriminator less, and should do so because they perceive the discriminator to have been less obligated to avoid discriminating. This research suggests that people's ideologies inform their judgments of obligations for implicit racial bias. The results also show that people blame those who violate obligations, suggesting the need for more research on obligation in moral and legal psychology. Together, these studies suggest that obligation judgments may mediate ideology–blame relationships. For transgressions that support or challenge social hierarchies, reasoning about obligations and blame may reflect preferences for hierarchical social structures.

**Acknowledgments** This research was supported by Project Implicit. Kate Ratliff is a consultant with Project Implicit, Inc., a nonprofit organization that includes in its mission “to develop and deliver methods for investigating and applying phenomena of implicit social cognition, including especially phenomena of implicit bias based on age, race, gender or other factors.”

### Compliance with Ethical Standards

**Conflict of interest** Liz Redford and Kate A. Ratliff declare that they have no conflict of interest.

## References

- Beierlein, C., Werner, C. S., Preiser, S., & Wermuth, S. (2011). Are just-world beliefs compatible with justifying inequality? Collective political efficacy as a moderator. *Social Justice Research, 24*(3), 278–296.
- Bertelsen, P. (2012). Intentional activity and free will as core concepts in criminal law and psychology. *Theory and Psychology, 22*(1), 46–66.

- Brady, J. B. (1996). Conscious negligence. *American Philosophical Quarterly*, 33(3), 325–335.
- Cameron, C. D., Payne, B. K., & Knobe, J. (2010). Do theories of implicit race bias change moral judgments? *Social Justice Research*, 23(4), 272–289.
- Carnes, M., Devine, P. G., Manwell, L. B., Byars-Winston, A., Fine, E., Ford, C. E., et al. (2015). Effect of an intervention to break the gender bias habit: A cluster randomized, controlled trial. *Academic Medicine*, 90(2), 221–230.
- Correll, J., Park, B., Judd, C., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the thin blue line: Police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology*, 92, 1006–1023.
- Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in Moral judgment: Testing three principles of harm. *Psychological Science*, 17(12), 1082–1089.
- Dasgupta, N. (2013). Implicit attitudes and beliefs adapt to situations: A decade of research on the malleability of implicit prejudice, stereotypes, and the self-concept. In P. G. Devine & E. A. Plant (Eds.), *Advances in Experimental Social Psychology* (Vol. 47, pp. 233–279). Cambridge: Academic Press.
- Duckitt, J. (2009). Punishment attitudes: Their social and psychological bases. In M. E. Oswald, S. Bieneck & J. Hupfeld-Heinemann (Eds.), *Social Psychology of Punishment of Crime* (pp. 137–156). Hoboken, NJ: Wiley.
- Duckitt, J., & Sibley, C. G. (2007). Right wing authoritarianism, social dominance orientation and the dimensions of generalized prejudice. *European Journal of Personality*, 21(2), 113–130.
- Eysenck, H. J. (1971). Social attitudes and social class. *British Journal of Social and Clinical Psychology*, 10, 210–221.
- Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., & Banaji, M. R. (2007). Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *Journal of General Internal Medicine*, 22(9), 1231–1238.
- Implicit Association Test: Are You Secretly Racist? (Hint: You Are) (2012, August 6). *Science 2.0*. Retrieved from [http://www.science20.com/news\\_articles/implicit\\_association\\_test\\_are\\_you\\_secretly\\_racist\\_hint\\_you\\_are-92783](http://www.science20.com/news_articles/implicit_association_test_are_you_secretly_racist_hint_you_are-92783).
- Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, 129(3), 339.
- Jost, J. T., & Hunyady, O. (2005). Antecedents and consequences of system-justifying ideologies. *Current Directions in Psychological Science*, 14(5), 260–265.
- Jost, J. T., Rudman, L., Blair, I. V., Carney, D. R., Dasgupta, N., Glaser, J., & Hardin, C. (2009). The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections and executive summary of ten studies that no manager should ignore. *Research in Organizational Behavior*, 29, 39–69.
- Kaiser, C. R., Dyrenforth, P. S., & Hagiwara, N. (2006). Why are attributions to discrimination interpersonally costly? A test of system-and group-justifying motivations. *Personality and Social Psychology Bulletin*, 32(11), 1423–1536.
- Kay, A. C., & Jost, J. T. (2003). Complementary justice: Effects of “poor but happy” and “poor but honest” stereotype exemplars on system justification and implicit

- activation of the justice motive. *Journal of Personality and Social Psychology*, 85(5), 823.
- Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The influence of intentionality and foreseeability. *Cognition*, 108, 754–770.
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. L., Joy-Gaba, J. A., et al. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143, 1765–1785.
- Levesque v. Anchor Motor Freight, Inc., 832 F.2d 702, 704 (1st Cir. 1987).
- Lipkus, I. M., Dalbert, C., & Siegler, I. C. (1996). The importance of distinguishing the belief in a just world for self versus for others: Implications for psychological well-being. *Personality and Social Psychology Bulletin*, 22(7), 666–677.
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2012). Moral, cognitive, and social: The nature of blame. In J. Forgas, K. Fiedler, & C. Sedikides (Eds.), *Social thinking and interpersonal behavior* (pp. 313–331). Philadelphia: Psychology Press.
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, 25(2), 147–186.
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, 37(5), 435–442.
- Merritt, R. D., & Harrison, T. W. (2006). Gender and ethnicity attributions to a gender- and ethnicity unspecified individual: Is there a people = White Male Bias? *Sex Roles*, 54(11–12), 787–797.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting intergroup attitudes and stereotypes from a demonstration website. *Group Dynamics*, 6, 101–115.
- Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ratliff, K. A., et al. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, 18(1), 36–88.
- O'Brien, L. T., & Major, B. (2005). System-justifying beliefs and psychological well-being: The roles of group status and identity. *Personality and Social Psychology Bulletin*, 31(12), 1718–1729.
- O'Brien, L. T., Crandall, C. S., Horstman-Reser, A., Warner, R., Alsbrooks, A., & Blodorn, A. (2010). But I'm no bigot: How prejudiced White Americans maintain unprejudiced self-images. *Journal of Applied Social Psychology*, 40(4), 917–946.
- Oldmeadow, J., & Fiske, S. T. (2007). System-justifying ideologies moderate status competence stereotypes: Roles for belief in a just world and social dominance orientation. *European Journal of Social Psychology*, 37(6), 1135–1148.
- Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67, 741–763.
- Redford, L., & Ratliff, K. A. (2015). Perceived moral responsibility for attitude-based discrimination. *British Journal of Social Psychology* (in press).
- Restatement (Second) of Torts, § 314A, 1977.
- Ridgeway, C. I., & Diekema, D. (1992). Are gender differences status difference? In C. I. Ridgeway (Ed.), *Gender, interaction, and inequality* (pp. 157–180). Berlin: Springer.
- Sekaquaptewa, D., & Espinoza, P. (2004). Biased processing of stereotype-incongruity

- is greater for low than high status groups. *Journal of Experimental Social Psychology*, 40, 128–135.
- Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.
- Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, 4, 267–281.
- Skitka, L. J. (2014). The psychological foundations of moral conviction. In J. Wright, & H. Sarkissian (Eds.), *Advances in Experimental Moral Psychology* (pp. 148–166). New York, NY: Bloomsbury Academic Press.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn and Bacon.
- Thomason v. Willingham, 165 S.E.2d 865, 867 (Ga. Ct. App. 1968).
- Thomsen, L., Green, E. G., & Sidanius, J. (2008). We will hunt them down: How social dominance orientation and right-wing authoritarianism fuel ethnic persecution of immigrants in fundamentally different ways. *Journal of Experimental Social Psychology*, 44(6), 1455–1464.
- Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, 10(1), 72–81.
- Uhlmann, E. L., Zhu, L., & Diermeier, D. (2014). When actions speak volumes: The role of inferences about moral character in outrage over racial bigotry. *European Journal of Social Psychology*, 44(1), 23–29.
- U.S. v. Carroll Towing, 159 F.2d 169 (2d Cir. 1947).
- Wilson, G. (2013). *The psychology of conservatism*. London: Routledge.