

# A cleansing fire: Moral outrage alleviates guilt and buffers threats to one's moral identity

Zachary K. Rothschild<sup>1</sup> · Lucas A. Keefer<sup>2</sup>

© Springer Science+Business Media New York 2017

**Abstract** Why do people express moral outrage? While this sentiment often stems from a perceived violation of some moral principle, we test the counter-intuitive possibility that moral outrage at third-party transgressions is sometimes a means of reducing guilt over *one's own* moral failings and restoring a moral identity. We tested this guilt-driven account of outrage in five studies examining outrage at corporate labor exploitation and environmental destruction. Study 1 showed that personal guilt uniquely predicted moral outrage at corporate harm-doing and support for retributive punishment. Ingroup (vs. outgroup) wrongdoing elicited outrage at corporations through increased guilt, while the opportunity to express outrage reduced guilt (Study 2) and restored perceived personal morality (Study 3). Study 4 tested whether effects were due merely to downward social comparison and Study 5 showed that guilt-driven outrage was attenuated by an affirmation of moral identity in an unrelated context.

**Keywords** Moral outrage · Guilt · Moral psychology · Self and identity · Defensive processes

## Introduction

When reports of the working conditions at Foxconn facilities in China revealed nightmarish conditions amounting to forced labor, many Americans expressed anger toward

companies (e.g. Apple) that had profited from the exploitation of Chinese laborers (Gupta 2012). In a word, people were *outraged* despite the fact that Americans themselves were not working in these facilities, and indeed many had purchased iPhones and other products that promoted these conditions.

Social psychologists use the term *moral outrage* to describe these feelings of anger, which are directed at a third-party for violating some moral standard of justice or fairness (Haidt 2003; Hoffman 2000; Leach et al. 2002; Montada and Schneider 1989; Vidmar 2000). Because this outrage is often expressed on behalf of the victim of that moral violation, moral outrage has been described as a prosocial emotion reflecting a desire to restore justice by fighting on behalf of the victimized (Thomas et al. 2009). Lauded as a force for promoting positive social outcomes, moral outrage is associated with numerous behaviors including support for political action (Montada and Schneider 1989; Thomas 2005), protest participation (Lodewijkz et al. 2008), and a desire to punish moral transgressors on behalf of innocent victims (e.g., Pagano and Huo 2007).

However, this altruistic portrayal of outrage has been called into question by research (reviewed below) showing that people express greater moral outrage following threats to their ingroup's moral status. These findings suggest that moral outrage is sometimes defensively grounded in a threat to one's own moral standing, rather than a desire for justice *per se*. Building on this premise, we present five studies testing whether moral outrage may sometimes be motivated by an effort to assuage personal guilt by casting aspersions on another.

---

✉ Zachary K. Rothschild  
zrothsch@bowdoin.edu

<sup>1</sup> Department of Psychology, Bowdoin College, 6900 College Station, Brunswick, ME 04011, USA

<sup>2</sup> University of Southern Mississippi, Hattiesburg, USA

## Moral identity concerns

Research demonstrates that people are motivated to see themselves and their social groups as acting in line with moral beliefs (Jordan et al. 2011; Mazar et al. 2008; Monin and Jordan 2009; Nisan 1991). In fact, people tend to view their ingroup's morality as *more* important than either its competence or sociability (Leach et al. 2007). Given that importance, it is perhaps not surprising that violations of moral standards elicit distinct responses aimed at restoring morality.

One such response is *guilt*, a self-focused, core "moral emotion" (Tangney et al. 2007) elicited by a violation of an internalized moral standard committed by oneself or one's group (O'Connor 2010; Branscombe et al. 2004; Tangney 1995). Guilt is a common response to perceived personal or in-group responsibility for harm perpetrated against an innocent victim and is therefore considered an essentially *social* emotion (Baumeister et al. 1994). In contexts ranging from a forgotten meeting all the way to racial genocide, *guilt* serves as a marker of a strained social bond in need of repair. Importantly, guilt relays to the individual that the bond has been damaged by *one's own actions*, and guilt is therefore a statement about the individual's own lack of moral standing.

A litany of research shows that guilt over personal or collective harm-doing often motivates guilt-reduction strategies aimed at repairing, undoing, or apologizing for perpetrated harm (Branscombe et al. 2002; Hoffman 2000; Iyer et al. 2003; McGarty et al. 2005; Stewart et al. 2010). For example, White Americans' guilt over salient racial inequality motivated support for reparative action to undo unjust status differences with Black Americans (Harvey and Oswald 2000).

However, given the potentially high cost of undoing harm, people often opt to engage in indirect methods of evading the pangs of their conscience (e.g., Branscombe and Miron 2004; Miron et al. 2010). For example, Tarrent et al. (2012) found that British and U.S. nationals who read accounts of their own group (vs. another group) perpetrating torture were more likely to say that torture was morally justified, presumably as a means of defusing guilt and protecting their group's moral identity.

Another psychological strategy for maintaining a positive moral identity is attributing blame to a focal target in order to exculpate perceived personal or collective responsibility. Rothschild et al. (2012) found that reminders of participants' own environmentally destructive behavior increased their willingness to blame corporations for harming the environment. This effect was due specifically to increased feelings of guilt over participants' own negative environmental impact. A follow-up study found that affirming participants' moral identity in an unrelated context

eliminated this guilt-driven blame. These findings are consistent with the idea that ascribing blame to a third-party can be motivated by a desire to defend the self against feelings of moral guilt and maintain a sense of moral identity without the often extensive cost of attempting to undo one's wrongdoing.

## Defensive outrage

Given that guilt can motivate people to perceive greater third-party responsibility for harm, it stands to reason that such threats may also elicit feelings of moral outrage. Consistent with this idea, Täuber and van Zomeren (2013) found that participants who experienced a threat to their group's moral (vs. nonmoral) status showed increased outrage at an outgroup target. This increase in outrage was explained as a defensive shift in participants' emotional focus from ingroup to outgroup that was spurred by a perceived threat to the ingroup's moral identity.

Rothschild and colleagues (2013) examined the potential mechanism behind threat-induced outrage by manipulating both the perceived cause of a disadvantaged group's suffering and the salience of a blameable third-party. They found that self-identified middle-class Americans made to feel their group was responsible for working-class suffering expressed increased moral outrage when given the opportunity to view illegal immigrants as a third-party harm-doer. Furthermore, elevated outrage elicited by this moral identity threat was associated with a perception that immigrants harm the working-class. However, in the absence of moral identity threat, perceptions of immigrant harm-doing were not associated with feelings of outrage. This finding suggests that the observed increase in outrage following a moral identity threat was a response to that threat specifically, and not merely the result of incidentally seeing immigrants as a threat to the working class.

## Present research

Feelings of guilt are a direct threat to one's sense that they are a moral person and, accordingly, research on guilt finds that this emotion elicits strategies aimed at alleviating guilt that do not always involve undoing one's actions. Furthermore, research shows that individuals respond to reminders of their group's moral culpability with feelings of outrage at third-party harm-doing.

These findings suggest that feelings of moral outrage, long thought to be grounded *solely* in concerns with maintaining justice may sometimes reflect efforts to maintain a moral identity. In other words, while outrage often aims to correct an injustice in the world, in some situations expressions of defensive outrage may be driven by a need to restore a moral identity. Past research showing that ingroup

harm-doing can elicit expressions of outgroup-directed outrage suggests that outrage may serve this function, however, this literature lacks any critical evidence that outrage actually provides any moral benefit.

The present research was designed to extend this literature by testing whether expressions of moral outrage can stem from a defensive motive to maintain a moral identity. We assessed evidence of this process by testing (1) whether outrage is elicited by a threat to one's moral identity specifically, (2) whether this effect is mediated and moderated by markers of perceived moral standing, and (3) whether such elicited expressions of outrage function to alleviate threat. The present studies test whether expressions of outrage against third-party harm-doing are sometimes spurred by concerns over one's moral standing rather than other, incidental negative feelings associated with guilt. We expect that this strategy will be effective, that is, outrage in response to a threatened moral identity will subsequently serve to attenuate guilt and restore one's perceived personal morality. In an effort to provide a broad demonstration of this phenomenon's breadth and importance we test this account across five studies in the context of moral outrage at corporate harm-doing for labor exploitation (Studies 1, 4, & 5) and harmful environmental conditions (Studies 2 & 3).<sup>1</sup>

## Study 1

Study 1 provided an initial test of the association between guilt and moral outrage when one's own immorality is salient. Specifically, we examined whether guilt following a reminder of personal culpability for sweatshop labor is associated with outrage at, and a resulting desire to punish, corporations for that exploitation.

This study builds on two key insights from prior research. First, that guilt motivates individuals to focus blame on a third-party (Rothschild et al. 2012) and further that moral identity threat elicits moral outrage and a desire to punish a transgressor (Rothschild et al. 2013). Based on the premise

<sup>1</sup> Given an issue can be framed in either moral or non-moral terms (Van Bavel et al. 2012) a pilot study was conducted to ensure that the presentation of the issues in our studies (sweatshop labor in Studies 1, 4 and 5; environmentally destructive behavior in Studies 2 and 3), was perceived in moral terms. A separate sample of 151 MTurk participants indicated the extent to which their attitudes about sweatshop labor and people's environmentally destructive behavior were "a reflection of your core moral beliefs and convictions" after reading the articles used in the primary studies. The item was adapted from previous research (Luttrell et al. 2016) and responses were made on a 7-point scale (1 = *not at all*, 7 = *extremely*). Supporting our assumption, one-sample *t* tests revealed that the means ( $M_{\text{sweatshop}} = 5.66$ ,  $M_{\text{environment}} = 5.03$ ) were significantly higher than the scales' midpoint (4),  $t(150) = 14.35$ ,  $p < .001$  and  $t(150) = 8.14$ ,  $p < .001$ , respectively.

that outrage at third-party harm-doing can be driven by guilt over one's own actions we hypothesized that:

**H1:** *Guilt will predict increased support for punishing third-party harm-doers, indirectly through increased moral outrage at third-party harm-doing.*

We also tested our claim that outrage and support for punishing third-party harm-doing would be due specifically to feelings of guilt and not general negativity by assessing and statistically controlling for general negative affect in our analyses.

## Method

Two-hundred and seventy-four American adults were recruited to participate through Amazon's Mechanical Turk (Mturk) service for \$.50.<sup>2</sup> Data from 40 cases was excluded from analyses *a priori* due to failing a key attention check (34 participants), reporting difficulty viewing study materials<sup>3</sup> (7 participants), and expressing suspicions (3 participants) during a final probing period. The remaining 234 participants (101 women) ranged in age from 18 to 71 years ( $M = 35.78$ ,  $SD = 12.04$ ). The average self-reported political orientation (1 = *very conservative*, 4 = *moderate*, 7 = *very liberal*) of our sample was skewed slightly liberal ( $M = 4.48$ ,  $SD = 1.77$ ). We initially conducted analyses for all five studies statistically controlling for political orientation. However, because political orientation had no effect on any of our hypothesized results we chose excluded this variable from our analyses for ease of presentation. It took participants on average 9.95 min to complete all survey materials.

As with all our studies, this experiment was presented as an examination of personality and attitudes about issues in the news that involved reading short news excerpts and completing personality and attitude questionnaires. This study employed a correlational design whereby all participants were exposed to the same materials, which are described below in the order that they were presented.

<sup>2</sup> To ensure sufficient power for the planned mediation analysis, we collected a large sample. In a simple mediation analysis (i.e., with a single mediator), and two small-to-moderate paths, a sample of 148 is necessary to achieve sufficient power (Fritz and McKinnon 2007). Given that the sequential mediation model incorporates two separate simple models, we collected approximately double this sample to ensure more than sufficient power for the planned analysis.

<sup>3</sup> Pilot studies revealed that a small number of participants were unable to properly view the news articles due to technical issues (e.g., slow internet speeds, hardware issues). Based on this, all participants were asked if they had any difficulty viewing the news articles upon completion of the studies and excluded accordingly.

## Materials and procedure

First, participants read a (fabricated) news article from a well-known newspaper outlet entitled “The Exploitation of Workers: A Blight on the Developing World.” The article discussed “Subhuman working conditions” across the developing world, including an “estimated 3 billion people and 250 million children working in so-called sweatshops, characterized by forced labor, substandard pay and hazardous working conditions.” The article accented the suffering of sweatshop workers by detailed examples of harmful sweatshop labor practices. Participants then rated the extent to which they believed “workers in developing countries were suffering as a result of sweatshop labor conditions” (1 = *not at all*, 7 = *very much*). A one-sampled *t* test revealed that participants responses to this item ( $M = 6.47$ ,  $SD = .82$ ) were significantly higher than scale’s midpoint (4),  $t(233) = 45.84$ ,  $p < .001$ , indicating a general acknowledgment that sweatshop workers experience high levels of harm.

### Personal responsibility induction

Afterwards, participants indicated the extent to which they personally engage in five behaviors purported to “directly or indirectly contribute to the perpetuation of sweatshops and forced child labor in the developing world.” In an effort to highlight participants’ personal culpability and thereby threaten their moral identity, we intentionally selected five behaviors assumed to be common for participants in our sample (“I buy products without knowing where they were made.”; “I rarely ask about working conditions when making a purchase”; “I sometimes buy products at big box stores such as Walmart, Target, K-mart, Sears etc.”; “When making a purchase I pay more attention to the cost than where it was made” ; “When making a purchase I don’t think about the workers who made the product I am purchasing.”) Responses were made on a 7-point scale (1 = *not at all true for me*, 7 = *very true for me*). Supporting our assumption that the items referred to common behaviors, a one-sample *t* test revealed that the grand mean of the composite scores averaging across the five items ( $M_{\text{grand}} = 5.81$ ,  $SD = 1.13$ ;  $\alpha = .86$ ) was significantly higher than the scale’s midpoint (4),  $t(233) = 24.45$ ,  $p < .001$ . The composite of participants’ responses to these five behavioral items served as an index of behavior contributing to labor exploitation.<sup>4</sup>

<sup>4</sup> Although participants generally admitted to engaging in behaviors contributing to labor exploitation, self-reported contribution scores were not correlated with our primary variables of interest. This is consistent with the idea that felt guilt, rather than the mere recognition of one’s harm-doing, is the driving force behind subsequent increases in moral outrage and retributive action against perceived corporate harm-doing.

### Guilt measure

Participants completed the Positive and Negative Affect Schedule (PANAS; Watson et al. 1988), a common measure of mood. Participants indicated the degree to which they were currently experiencing 10 positive emotions ( $M_{\text{grand}} = 2.76$ ,  $SD = .87$ ;  $\alpha = .91$ ) and 10 negative emotions ( $M_{\text{grand}} = 1.61$ ,  $SD = .29$ ;  $\alpha = .90$ ) using a 5-point response scale (1 = *very slightly or not at all*, 5 = *extremely*). Of particular interest for the current study, participants indicated the extent to which they felt “Guilty.” Responses to this item comprised our measure of personal guilt ( $M = 1.97$ ,  $SD = 1.08$ ). The composite of the 9 remaining negative affect items was also calculated to test the specificity of guilt’s effects ( $M = 1.57$ ,  $SD = .64$ ).

### Moral outrage measure

Participants completed a 6-item moral outrage measure (used in Pagano and Huo 2007; Rothschild et al. 2013) to assess anger at a third-party for perpetrating harm against a victimized outgroup. Specifically, participants indicated the degree to which they felt anger at *international corporations* for the harm done to sweatshop workers in the developing world (e.g., “Thinking about the injustices that some workers in developing countries may have suffered from international corporations’ abusive labor practices makes me angry.”; “Knowing that workers in developing countries are probably helpless against corporations’ abusive labor practices makes me angry on their behalf.”). Responses were made on a 7-point scale (1 = *not at all*, 7 = *very much*) and were averaged to form composite scores ( $M_{\text{grand}} = 5.54$ ,  $SD = 1.51$ ;  $\alpha = .97$ ).

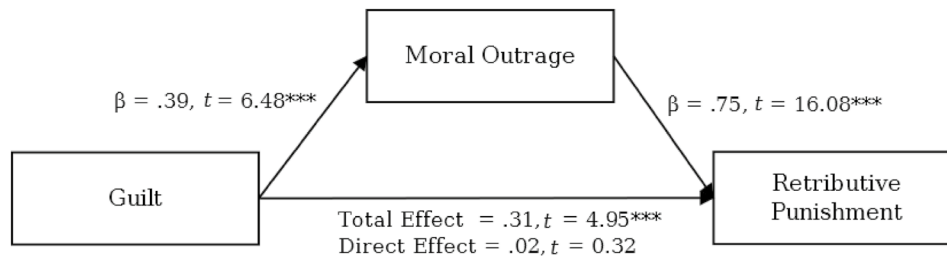
### Support for retributive punishment

Finally, participants completed a 4-item measure of retributive punishment used in previous research to assess the desire to punish a third-party perpetrator for harming a victimized outgroup (Rothschild et al. 2013). Specifically, participants indicated the degree to which they support greater efforts to punish International Corporations for the harm perpetrated against sweatshop workers in the developing world (e.g., “International Corporations should face harsher punishment for the harm they cause workers in developing countries” ; “Whatever the cost, corporations must be brought to justice for unjustly hurting workers in developing countries.”). In an effort to more directly measure participants’ willingness to exact punishment on corporate harm-doers we added an addition item asking participant if they, “would see that corporations were severely punished for exploiting workers” if they personally had the power to do so. Responses for all items were made on a

**Table 1** Observed correlations between the variables observed (Study 1)

	Guilt	Moral outrage	Retributive punishment	Negative affect
Guilt	–	.39***	.31***	.61***
Moral outrage		–	.76***	.22***
Retributive punishment			–	.18**
Negative affect				–

Note \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$



Total adjusted  $R^2$  for the model = .57,  $F(2, 231) = 155.03, p < .001$ .

**Fig. 1** Mediation of the effect of guilt on retributive punishment through moral outrage (Study 1). Note: All path coefficients represent standardized regression weights. The direct effect coefficient repre-

sents the effect of guilt on retributive punishment after controlling for the effect of moral outrage. \*Significant at  $p < .05$ ; \*\*Significant at  $p < .01$ ; \*\*\*Significant at  $p < .001$

7-point scale (1 = not at all, 7 = very much) and were averaged across all five items to form composite scores ( $M_{grand} = 5.58, SD = 1.42; \alpha = .95$ ).

**Results**

We initially tested the correlations among our three variables of interest: guilt, moral outrage, and retributive punishment. The full table of correlations is present in Table 1 with and without controlling for general negative affect. Consistent with our expected model, we found that guilt, moral outrage, and support for retributive punishment were each significantly and positively correlated with one another.

Next we tested the hypothesized mediation model in which increased guilt predicts support for retributive punishment through increased moral outrage. We did this by conducting an indirect effects analysis with guilt as the primary predictor, third-party punishment as the outcome, and moral outrage as the proposed mediator with general negative affect included as a covariate (for full model information, see Fig. 1). Bootstrapping analysis with 5000 resamples returned a 95% confidence interval for the indirect effect of (.27, .56); providing evidence at  $\alpha < .05$  that guilt increased punitiveness indirectly through an increase in third-party-directed outrage. The indirect effect remained

significant even when general negative affect was not included as a control (95% CI: .28, .51).

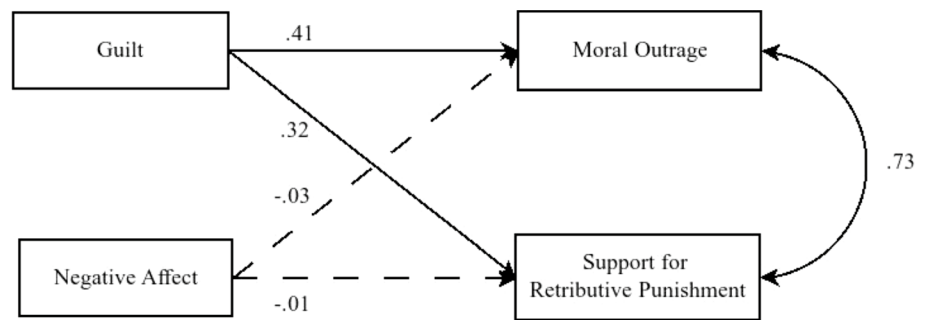
**Post hoc model comparisons**

While this mediation analysis confirmed our initial predictions about the associations between the variables, the correlational design invites questions about the most appropriate way to model their relationships. In particular, it is unclear whether participants expressed support for retributive punishment as a function of outrage or, alternatively, whether support for punishment motivated the expression of outrage as a justification for this support (i.e., the reverse model). Therefore, we conducted a series of post hoc model comparisons to determine whether our purported account of the process (in the mediation model above) is a better fit to the data. We began with a simple path model treating outrage and support for punishment as two associated outcomes (see Fig. 2 for summary).

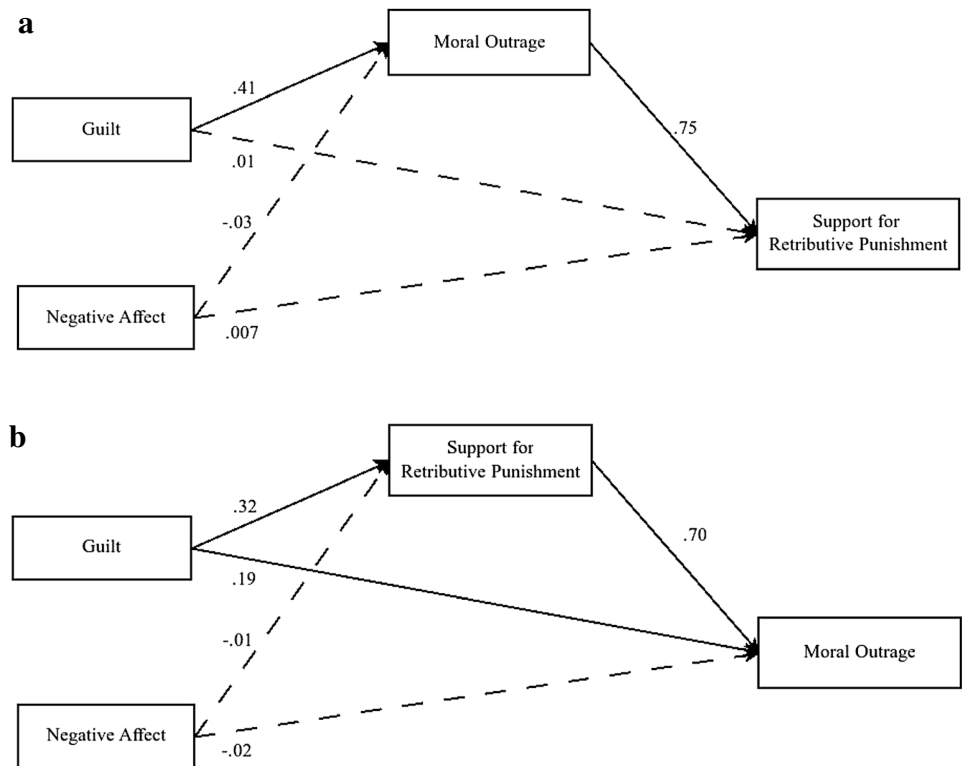
We then compared two possible variations of this model: One in which outrage is treated as a mediator (Fig. 3a) and another in which it is treated as the outcome (Fig. 3b). While these models fit the data just as well as the initial model (all contain the same number of parameters), we found that outrage effectively accounted for the effect of guilt on retributive punishment but not vice versa. Further confirming our proposed account, we found that simplifying this model by eliminating the non-significant direct



**Fig. 2** Baseline model with two correlated outcomes (Study 1). Note. Path estimates are standardized. Dashed paths are non-significant and all solid paths significant at  $p < .001$



**Fig. 3 a** Proposed model of the associations between variables (Study 1). Note. Path estimates are standardized. Dashed paths are non-significant and all solid paths significant at  $p < .001$ . **b** Alternative model of the associations between variables (Study 1). Note. Path estimates are standardized. Dashed paths are non-significant and all solid paths significant at  $p < .001$



effects of guilt and negative affect on retributive punishment (Fig. 4a) resulted in no appreciable loss of model fit ( $\chi^2(2) = .12, p = .94$ ). In other words, this more parsimonious model in which outrage fully explains the effect of guilt on punishment accurately represents our data. In contrast, eliminating the direct effects of affect on outrage in the alternative model (Fig. 4b) caused a marked lack of fit ( $\chi^2(2) = 15.64, p = .0004$ ), and therefore is a poorer representation of the data.

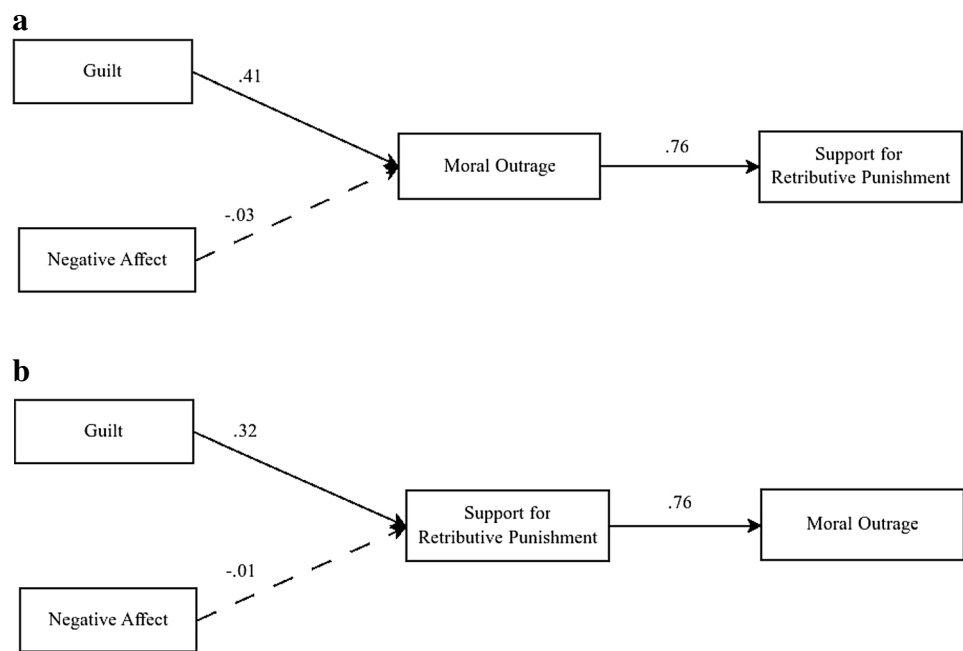
### Discussion

The results of Study 1 showed that guilt predicted increased punitiveness toward a third-party harm-doer due to increased moral outrage at the target. These results provide

initial support for the proposed guilt-driven account of moral outrage. Specifically, they support the view that guilt, a marker of perceived moral identity threat, is associated with greater outrage and punishment of third-party harm-doing, despite the common belief that these emotional and behavioral responses are independent of egoistic concerns. The fact that these effects emerged despite controlling for general negative affect suggests that the relationship between and guilt and outrage is unlikely to be the result of non-moral negative affect: Rather, outrage and a desire for punishment appear to be motivated by guilt specifically.

Although these findings are consistent with the hypothesized account of guilt-driven outrage, there are a few notable limitations that should be addressed. Most importantly, the correlational design employed in Study 1 limits our ability to draw firm causal conclusions. For instance,

**Fig. 4** **a** Simplified mediational model of the associations between variables (*Study 1*). *Note.* Path estimates are standardized. *Dashed paths* are non-significant and all *solid paths* significant at  $p < .001$ . **b** Alternative mediational model of the associations between variables (*Study 1*). *Note.* Path estimates are standardized. *Dashed paths* are non-significant and all *solid paths* significant at  $p < .001$



outrage at perceived corporate harm-doing may have incidentally increased feelings of guilt rather than the reverse. Furthermore, it is possible that the observed association between guilt and outrage was an artifact of some third variable, such as differences in the perceived importance of labor issues. Reminders of one's role in supporting sweatshop labor may have independently increased guilt and moral outrage. Given the symbiotic relationship between corporations and consumer behavior, this explanation is not unreasonable. An experimental design is needed to rule out these possibilities.

## Study 2

Study 2 was designed to replicate and extend the finding that personal guilt predicts moral outrage at a third-party. We did this by assessing participants' guilt over their own negative environmental impact before assessing their outrage at environmental destruction caused by oil companies.

By employing an experimental design, we also sought to directly test the causal role of guilt. First, we manipulated perceived culpability for the environmentally destructive effects of climate change by suggesting to participants that their ingroup (American consumers) or an outgroup (Chinese consumers) is responsible. We then manipulated whether participants reported guilt *before* or *after* having the opportunity to express moral outrage at corporate harm-doing. We hypothesized that:

**H2:** *A threat to one's moral identity in the form of salient ingroup (vs. outgroup) responsibility for harm will elicit increased outrage at third-party harm-doing indirectly through felt guilt prior to, but not after, expressing outrage.*

A third goal of Study 2 was to test the hypothesized defensive function of moral outrage. Based on the premise that third-party-directed outrage serves to assuage underlying feelings of guilt induced by a threat to one's own perceived morality, we hypothesized that:

**H3:** *Following the moral identity threat of salient ingroup (vs. outgroup) responsibility for harm, guilt assessed after expressing third-party-directed outrage will be lower than guilt assessed pre-outrage. Furthermore, ingroup (vs. outgroup) responsibility for harm-doing will increase guilt assessed prior to expressions of third-party-directed outrage, but have no effect on post-outrage guilt.*

Given the harmful nature of environmental destruction, we considered the possibility that participants might perceive themselves as victims, or potential victims of environmental destruction. Perceiving oneself to be the victim of illegitimate harm or insult has been shown to elicit feelings of *personal anger*, an emotion that is similar, yet conceptually distinct from moral outrage (Thomas et al. 2009). This raises the possibility that expressions of outrage at environmentally harmful corporate practices may actually reflect feelings of personal anger in response to participants' own perceived victimization. We address this possibility by

assessing, and statistically controlling for, the perceived victimization posed by environmental destruction.

## Method

Three-hundred and seven American adults were recruited to participate on Mturk for \$.75.<sup>5</sup> Data from 40 cases were excluded from analyses *a priori* due to either failing a key attention check (21 participants), reporting difficulty viewing study materials (17 participants), and/or completing survey materials too quickly (less than 5 min; 6 participants). The remaining 267 participants (148 women) ranged in age from 18 to 67 years ( $M=31.79$ ,  $SD=10.23$ ). It took participants on average 13.97 min to complete all survey materials. The average self-reported political orientation (1 = *very conservative*, 4 = *moderate*, 7 = *very liberal*) of our sample was skewed slightly liberal ( $M=4.51$ ,  $SD=1.53$ ). Participants were randomly assigned in a 2 (*Responsibility: Ingroup vs. Outgroup*) $\times$ 2 (*Order: Pre-outrage vs. Post-outrage* guilt assessment) between-subjects design.

## Materials and procedure

### Responsibility manipulation

All participants first read an article, ostensibly written by scientific experts, entitled “The Causes and Consequences of Climate Change Today”. The first section of this fabricated article was identical across conditions and discussed the negative effects of “man-made” climate change, including animal extinctions and the destructive weather events. The second section, entitled “Who’s to Blame?” differed between conditions.

In the *ingroup responsible* condition, this section was designed to threaten participants’ moral identity by identifying American consumers, an ingroup for all participants, as the primary source of climate change, noting that “Americans drive more, consume more energy at home, and waste more than the rest of the world” (following Rothschild et al. 2012). The article concluded that until “Americans start to acknowledge their responsibility for the damaging effects of climate change, things are likely to only get worse.” In the *outgroup responsible* condition, this section of the article was identical except that Chinese consumers were identified as the primary source of climate change.

### Perceived victimization measure

Following that manipulation, participants rated their agreement with three statements assessing the extent to which they perceived themselves or close others as being victimized by climate change (“Climate change has a direct negative effect on my life.”; “I am a victim of climate change.”; “I have friends/or family that have been victimized by the harmful effects of climate change.”; 1 = *strongly disagree*, 5 = *strongly agree*). Responses were averaged to create a composite perceived victimization score ( $\alpha=.87$ ;  $M=3.62$ ,  $SD=1.62$ ). As noted above, perceived victimization scores were included as a covariate for all primary analyses.<sup>6</sup>

### Guilt measure

Participants completed a validated 3-item measure of personal guilt over a specific outcome (e.g., Rothschild et al. 2015). Specifically, participants indicated the extent (1 = *strongly disagree*, 7 = *strongly agree*) to which they agreed that they felt *guilty*, *regretful*, and *apologetic* for the negative impact their lifestyle has on the environment. Responses were averaged to form composite personal guilt scores ( $M_{\text{grand}} = 4.25$ ,  $SD = 1.73$ ;  $\alpha = .92$ ).

### Moral outrage measure

Participants completed the same 6-item measure of moral outrage used in Study 1, which was modified to assess anger at the environmental destruction caused by multinational oil corporations (e.g., “I can’t help but feel angry when I think about the damage done to the environment caused by multinational oil corporations.”;  $M_{\text{grand}} = 4.74$ ,  $SD = 1.79$ ;  $\alpha = .98$ ). Importantly, the responsibility manipulation did not reference to the role of corporations, allowing scores on this measure to reflect participants’ spontaneous expression of outrage at a (viable) third-party.

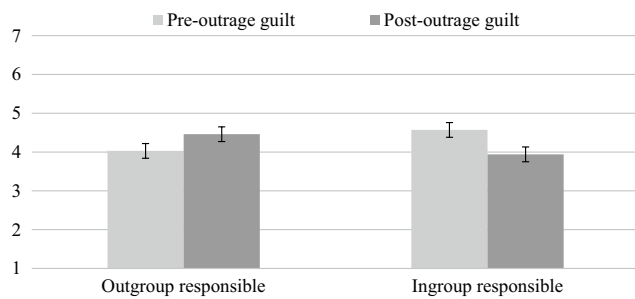
### Order manipulation

The order in which participants completed the personal guilt measure and moral outrage measure comprised our second experimental manipulation. Participants randomly assigned to the *pre-* (vs. *post-*) *outrage guilt* condition, completed the measure of personal guilt prior to (vs. after) expressing moral outrage at third-party perpetrated harm-doing.

<sup>5</sup> A sample size analysis using G\*Power (Faul et al. 2007) revealed that in order to ensure .80 power for our primary analysis, assuming a small to medium effect size we would need a total sample size of 264. Factoring in the exclusion rate from Study 1 we estimated that we would need to collect approximately 307 participants.

<sup>6</sup> A Responsibility $\times$ Order ANOVA on Perceived Victimization yielded no significant effects (*responsibility*:  $F(1, 263)=.04$ ,  $p=.85$ ; *order*:  $F(1, 263)=.07$ ,  $p=.80$ ; *responsibility* $\times$ *order*:  $F(1, 263)=.17$ ,  $p=.68$ ).





**Fig. 5** Guilt as a function of responsibility and order (*Study 2*). Black bars represent standard error

## Results

### Moral outrage

A 2 (*responsibility*) $\times$ 2 (*order*) ANCOVA on moral outrage at corporate harm-doing, controlling for perceived victimization scores yielded the predicted threat effect,  $F(1, 262)=8.82, p=.003, \eta_p^2=.03$ , such that participants in the *ingroup responsible* condition reported significantly higher levels of outrage at the environmental destruction caused by multinational oil companies ( $M=5.09, SD=1.55$ ) than those in the *outgroup responsible* condition ( $M=4.53, SD=1.84$ ). This effect remained significant when perceived victimization was not included as a covariate ( $F(1, 263)=7.47, p=.007, \eta_p^2=.03$ ). No other effects were significant (*order*:  $F(1, 262)=1.62, p=.20$ ; *responsibility* $\times$ *order*:  $F(262)=.90, p=.34$ ).

### Guilt

Submitting guilt scores to the same ANCOVA analysis yielded no significant main effects (*responsibility*:  $F(1, 262)=.003, p=.96, \eta_p^2<.001$ ; *order*:  $F(1, 262)=.30, p=.59, \eta_p^2=.001$ ), but did reveal the predicted responsibility $\times$ order interaction,  $F(1, 262)=8.69, p=.003, \eta_p^2=.03$  (See Fig. 5 for the pattern of means). This interaction remained significant when perceived victimization was not included as a covariate ( $F(1, 263)=7.59, p=.01$ ).

Pairwise comparisons (Fisher's LSD) revealed that *prior* to expressing moral outrage at third-party harm-doing, participants in the *ingroup responsible* condition, reported significantly higher guilt scores ( $M=4.57, SD=1.59$ ) than those in the *outgroup responsible* condition who were primed with outgroup culpability for climate change ( $M=4.03, SD=1.84; F=4.51, p=.04, \eta_p^2=.02$ ). In contrast, *after* having the opportunity to express outrage at climate change caused by third-party harm-doing, participants self-reported levels of guilt in the *ingroup responsible* condition ( $M=3.94, SD=1.66$ ) was significantly lower

than those the *outgroup responsible* condition ( $M=4.46, SD=1.76; F=4.18, p=.04, \eta_p^2=.02$ ).

Also consistent with predictions, for participants in the *ingroup responsible* condition, self-reported guilt scores were significantly lower when assessed after (vs. before) moral outrage ( $F=6.17, p=.01, \eta_p^2=.02$ ). In contrast, for participants in the *outgroup responsible* condition, guilt scores before and after expressing third-party-directed outrage were statistically equivalent ( $F=2.85, p=.10, \eta_p^2=.01$ ).

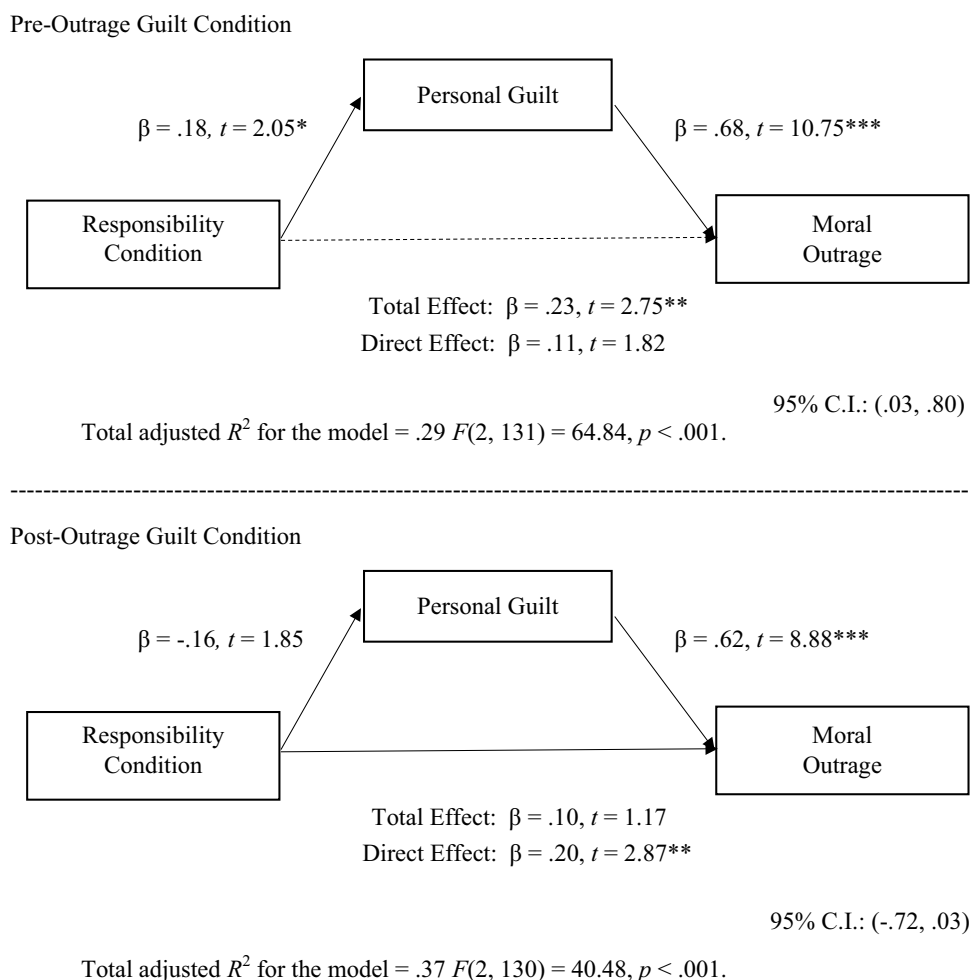
### Mediation analyses

We conducted a mediated moderation analysis using Preacher and Hayes' (2008) bootstrapping procedure. We first regressed moral outrage scores onto the interaction of responsibility condition (coded: *ingroup responsible*=1; *outgroup responsible*=0) and order (coded: *post-outrage guilt*=1; *pre-outrage guilt*=0) with self-reported guilt entered as a potential mediator and our main effects as covariates. The 95% confidence interval (5000 resamples) obtained for the indirect effect of the Responsibility $\times$ Order interaction on moral outrage did not contain zero [CI=-1.27, -.20]. We interpreted this finding by examining the conditional indirect effects for each *order* condition. As predicted, whereas the 95% confidence interval obtained for the indirect effect of responsibility on moral outrage through pre-outrage guilt did not contain zero [CI=.03, .80], the 95% confidence interval in the post-outrage guilt condition did contain zero [CI=-.72, .03] (See Fig. 6 for conditional indirect effects and parameters). These results suggests that variations in guilt assessed *before*, but not *after* expressing outrage at corporate harm-doing explain the effect of threat on moral outrage.

## Discussion

Supporting our first prediction, ingroup (vs. outgroup) culpability increased moral outrage at oil corporations' environmentally destructive practices. These threat-induced feelings of outrage were specifically due to increased feelings of personal guilt reported prior to, but not after, expressing outrage at third-party harm-doing. Importantly, these results, which are consistent with the proposed guilt-driven account of moral outrage, help rule out the possibility that threat-induced guilt was driven by elevated feelings of outrage (reverse causality), or that the relationship between guilt and outrage was an artifact of independent effects of perceived complicity in corporate harm-doing (third variable problem). Additionally, the fact that the effects on moral outrage remained when controlling for participants' perceived personal victimization suggests that

**Fig. 6** Conditional indirect effect of threat on moral outrage through guilt assessed *prior* to, and *after*, expressing outrage (Study 2). Note: All path coefficients represent standardized regression weights. The direct effect coefficient represents the effect of responsibility condition on moral outrage after controlling for the effect of personal guilt assessed prior to, or after, expressing outrage. \*Significant at  $p < .05$ ; \*\*Significant at  $p < .01$ ; \*\*\* Significant at  $p < .001$



the participants' outrage at corporate harm-doing did not merely reflect personal anger.

Supporting our second prediction, Study 2 found that the threat-induced increase in guilt over personal harm-doing was not present after expressing outrage at third-party harm-doing. In fact, among those given the opportunity to express outrage at corporate harm-doing, those initially primed with ingroup (vs. outgroup) culpability for environmental harm actually reported lower personal guilt. Also, supporting our predictions, participants threatened with ingroup harm-doing reported significantly less guilt after expressing outrage at corporate harm-doing. These findings are consistent with our general claim that expressions of third-party-directed outrage can represent a motivated defensive process aimed at protecting one's own perceived morality. Unexpectedly, we did find that the conditional effect of the responsibility manipulation on outrage became stronger after accounting for post-outrage guilt. This suggests that other motives for outrage (unrelated to guilt) were also present.

However, our ability to draw strong conclusions about the purported moral identity maintenance function of outrage is somewhat limited by Study 2's design. For one, the use of order as a manipulation in Study 2 allowed us to test the guilt-reducing effect of outrage but it leaves time as a potential confound. It is possible that decreased guilt after (vs. before) expressing outrage may have been due, not to this expression, but to differences in the temporal distance between the manipulation of responsibility and assessment of guilt. Furthermore, while guilt was assessed in Studies 1 and 2 as an emotional marker of a perceived moral identity threat, these studies did not directly assess the extent to which our responsibility manipulation was experienced as a threat to participants' personal moral identity. As noted, guilt also involves concern that one's immoral actions have potentially harmed, disappointed, or offended others (Baumeister et al. 1994). However, our claim is that outrage is a defense against the negative self-evaluation at the core of guilt rather than these peripheral anxieties about the social effects of immorality. Study 3 was designed to address these issues.

### Study 3

In Study 3 we directly assessed participants' perceived personal moral character after manipulating ingroup (vs. outgroup) culpability. Assessing personal moral identity as an outcome measure allowed us to both assess the validity of our responsibility manipulation and to directly test the proposed moral identity maintenance function of moral outrage.

We also addressed the confounding effect of a delay in Study 2. Unlike that study, only participants in the *third-party-directed outrage* condition were given the opportunity to express outrage at corporate harm-doing by completing a measure of moral outrage. Remaining participants assigned in the control condition instead completed an unrelated questionnaire of equal length. This manipulation allowed us to vary whether or not participants could express outrage while ensuring a roughly equivalent delay between responsibility manipulation and moral identity assessment.

Based on the premise that moral outrage serves to defend against the moral identity threat posed by personal or collective harm-doing, we predicted that:

**H4:** *The threat of ingroup (vs. outgroup) responsibility for harm-doing will reduce perceived personal moral character ratings, unless participants have the opportunity to express moral outrage at third-party harm-doing. Furthermore, in response to the threat of ingroup perpetrated harm, greater moral outrage at third-party harm-doing should predict greater personal moral character ratings.*

Importantly, for moral culpability, participants must believe that their own and others' actions can play a role in fostering the environmentally destructive effects of climate change. Thus, to ensure that our predicted effects were not merely due to between-condition differences in participants' belief in humanity's role in climate change, we assessed and statistically controlled for participants' belief in anthropogenic climate change in Study 3.

### Method

Three-hundred and seven Americans were recruited to participate on Mturk for \$.75.<sup>7</sup> Data from 61 cases were excluded from analyses *a priori* due to failing a key

attention check (37 participants), reported difficulty viewing study materials (28 participants), and/or completing survey materials too quickly (less than 5 min; 8 participants). The remaining 246 participants (155 women) ranged in age from 18 to 68 years ( $M=32.09$ ,  $SD=10.87$ ). It took participants on average 14.60 min to complete all survey materials. The average self-reported political orientation (1=*very conservative*, 4=*moderate*, 7=*very liberal*) of our sample was skewed slightly liberal ( $M=4.52$ ,  $SD=1.47$ ). Participants were randomly assigned in a 2 (*Responsibility: Ingroup responsible vs. Outgroup responsible*) $\times$ 2 (*Moral Outrage: Third-party-directed outrage vs. Control*) between-subjects design.

### Materials and procedure

Prior to any experimental manipulation, participants first responded to a single-item measure assessing belief in anthropogenic climate change ("I believe that human behavior contributes to climate change"; 1=*strongly disagree*, 5=*strongly agree*;  $M=4.17$ ,  $SD=.90$ ). A one-sampled *t* test revealed that participants' responses to this item were significantly higher than scale's midpoint (3),  $t(245)=20.49$ ,  $p<.001$ , indicating a general belief in anthropogenic climate change. Scores on this measure were used as a covariate for all analyses.

#### Responsibility manipulation

Using the procedure and materials described in Study 2, participants read a purportedly scientific article identifying either the participants' ingroup (Americans; *ingroup responsible*), or an outgroup (the Chinese; *outgroup responsible*) as the primary cause of climate change.

#### Moral outrage manipulation

Next participants were randomly provided one of two surveys as our second experimental manipulation. Participants assigned to the *third-party-directed outrage* condition were presented with the same 6-item moral outrage measure used in Study 2 ( $M_{grand} = 4.47$ ,  $SD=1.75$ ;  $\alpha = .97$ ). Completing this measure oriented participants towards a third-party harm-doer (i.e., multinational oil corporations) and gave them the opportunity to express moral outrage at them.

In contrast, participants assigned to the *control* condition were not given the opportunity to express outrage at third-party harm-doing. Instead, participants completed a purported personality survey instructing them to indicate how much physical pain they might experience (1=*not at all*, 7=*very much*) in six hypothetical scenarios (e.g. "You have dental surgery."; "You are in a car accident."; "You lose a

<sup>7</sup> A sample size analysis using G\*Power (Faul et al. 2007) revealed that in order to ensure .80 power for our primary analysis, assuming a small to medium effect size we would need a total sample size of 264. Factoring in the exclusion rate from the previous studies we estimated that we would need to collect approximately 307 participants.

limb.”;  $\alpha = .71$ ;  $M = 5.40$ ,  $SD = 1.01$ ). This potentially aversive control condition was used to decrease the likelihood that any potential effects might be explained by general negative affect rather than moral outrage specifically.

### Personal moral character measure

Finally, participants completed a purported personality survey (adapted from Zhong et al. 2010) instructing them to rank themselves in comparison to other people they know along several positive trait dimensions (moral character, sense of humor, creativity, fitness, social sensitivity, leadership). Responses were made in percentiles ranging from 0 (*worse than all others*) to 100 (*better than all others*). In line with previous research (Rothschild et al. 2015; Zhong et al. 2010), we used participants’ response to the moral character item as a measure of their perceived personal morality ( $M_{grand} = 74.32$ ,  $SD = 17.87$ ).

## Results

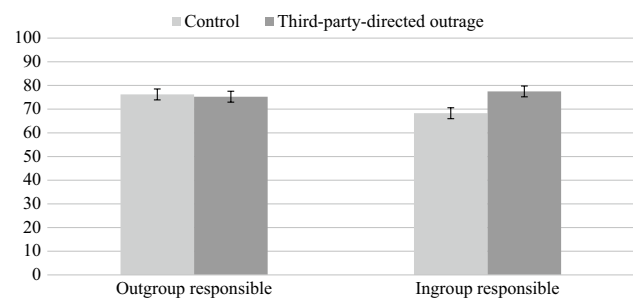
### Moral outrage

A one way ANCOVA on moral outrage at corporate harm-doing, controlling for belief in climate change yielded the predicted threat effect,  $F(1, 119) = 4.98$ ,  $p = .03$ ,  $\eta_p^2 = .04$ , such that participants in the *ingroup responsible* condition reported significantly higher levels of outrage at environmental harm caused by multinational oil companies ( $M = 4.78$ ,  $SD = 1.54$ ) than those in the *no threat* condition ( $M = 4.15$ ,  $SD = 1.91$ ). This effect was in the same direction, although marginal, when belief in climate change scores were not included as a covariate,  $F(1, 120) = 3.07$ ,  $p = .08$ .

### Personal moral character

A  $2$  (*responsibility*)  $\times$   $2$  (*third-party-directed outrage*) ANCOVA on perceived personal moral character, controlling for belief in anthropogenic climate change yielded no significant main effects (*responsibility*:  $F(1, 241) = 1.60$ ,  $p = .21$ ,  $\eta_p^2 = .007$ , *moral outrage*:  $F(1, 241) = 3.38$ ,  $p = .07$ ), but did reveal the predicted two-way interaction,  $F(1, 241) = 5.12$ ,  $p = .03$ ,  $\eta_p^2 = .02$  (see Fig. 7). This interaction effect remained significant when belief in anthropogenic climate change scores were not included as a covariate,  $F(1, 242) = 4.64$ ,  $p = .03$ . The interaction also remained significant when controlling for participants’ comparative rankings of their standing on morally irrelevant positive traits,  $F(1, 240) = 4.07$ ,  $p = .04$ .

As predicted, pairwise comparisons (Fisher’s LSD) revealed that in the absence of an opportunity to express



**Fig. 7** Personal moral character as a function of responsibility and opportunity to express third-party-directed outrage (*Study 3*). Black bars represent standard error

outrage at third-party harm-doing, those primed with ingroup culpability for harmful climate change rated themselves as having significantly lower personal moral character ( $M = 68.29$ ,  $SD = 19.32$ ) than those primed with outgroup culpability for climate change ( $M = 76.22$ ,  $SD = 16.85$ ;  $F = 6.25$ ,  $p = .01$ ,  $\eta_p^2 = .03$ ). In contrast, among those given the opportunity to express outrage at third-party harm-doing, participants personal moral character ratings in the *ingroup responsible* condition ( $M = 77.50$ ,  $SD = 15.66$ ) did not significantly differ from those the *outgroup responsible* condition ( $M = 75.26$ ,  $SD = 18.60$ ;  $F = .50$ ,  $p = .48$ ,  $\eta_p^2 = .002$ ).

Also consistent with predictions, for participants in the *ingroup responsible* condition, participants given the opportunity to direct moral outrage at third-party harm-doing reported significantly higher personal moral character ratings than those not given the opportunity to express moral outrage ( $F = 8.52$ ,  $p = .004$ ,  $\eta_p^2 = .03$ ). In contrast, for participants in the *outgroup responsible* condition, personal moral character ratings were equivalent regardless of whether or not the participants were able to express third-party-directed outrage ( $F = .09$ ,  $p = .77$ ,  $\eta_p^2 < .001$ ).<sup>8</sup>

## Discussion

First, supporting the validity of our responsibility manipulation as a moral identity threat, participants primed with ingroup (vs. outgroup) culpability for environmental

<sup>8</sup> We conducted partial correlational analyses to assess the relationship between moral outrage at third-party harm-doing and perceived personal moral character within conditions. This analysis revealed a significant positive correlation between moral outrage scores and personal moral character ratings following a moral identity threat,  $r(60) = .25$ ,  $p = .05$ . This correlation remained significant when belief in anthropogenic climate change scores were not included as a control,  $r(63) = .26$ ,  $p = .04$ . There was no significant correlation between outrage and moral character in the absence of a moral identity threat,  $r(56) = .03$ ,  $p = .85$ .



harm-doing reported lower personal moral character ratings, at least when there was no opportunity to express moral outrage. However, as predicted, this effect was attenuated when participants were given a chance to express moral outrage. Furthermore, consistent with the pattern of results from Study 2, participants exposed to moral identity threatening information reported greater outrage.

Building on the previous two studies, the results of Study 3 provide even more direct evidence for the proposed account of defensive outrage and its role as a moral identity maintenance strategy. Specifically, Study 3 shows that expressions of moral outrage can be motivated by threats to one's own moral identity and serve to bolster one's perceived personal morality. However, as in Study 2, our ability to draw firm conclusions about the defensive function of moral outrage is still limited by an important confound. Specifically, in manipulating whether or not participants could express moral outrage in Study 3, we also inadvertently manipulated whether or not third-party harm-doing was made salient to participants. Therefore, Study 3 confounds the opportunity to express moral outrage (vs. not) with the presence (vs. absence) of salient third-party transgressor. Previous research has shown that perceived threats to the self can motivate people to engage in downward social comparisons (comparing oneself with someone who is worse) to protect their own self-esteem (Wills 1981). Thus, participants who had the opportunity to express outrage in Study 3 also received a reminder of a group that is comparatively more harmful to the environment than themselves. Comparing themselves to this group may have minimized their own wrongdoing, completely independent of any expression of outrage.

This introduces the possibility that expressing third-party-directed outrage following a moral identity threat may have bolstered participants' personal morality (Study 3) and decreased their guilt (Study 2), not due to the expression of moral outrage *per se*, but due to participants' ability to trivialize their own actions in light of the comparably worse actions of corporations. If true, this suggests that increased feelings of moral outrage may arise merely from the tendency to see another's harm-doing as significantly worse than one's own. In order to rule out this alternative explanation it is necessary to separate participants' *perception* of third-party harm-doing from their *ability to express moral outrage*.

## Study 4

The primary goal of Study 4 was to test whether mere exposure to third-party harm-doing, rather than the ability to express outrage at third-party-harm-doing specifically, accounts for the findings observed in Study 3. As in Study 1, all participants were first reminded of harmful sweatshop

labor conditions in the developing world. However, participants then indicated the extent to which they blamed a third-party (international corporations) for the harmful effects of sweatshop labor. This ensured that all participants were given the opportunity to compare their role in promoting sweatshop labor to a third-party. As in Study 3, we then manipulated whether or not participants were given the opportunity to express outrage at corporate harm-doing. Finally, participants' rated their own perceived personal moral character. Based on the premise that expressing outrage at a third-party harm-doer serves a moral identity maintenance function that cannot be accounted for by downward social comparison alone, we hypothesized that:

**H5:** *Among those who blame a third-party perpetrator for illegitimate harm-doing, the opportunity to express (vs. not express) third-party-directed outrage will increase participants' personal moral character ratings.*

## Method

One-hundred and forty-eight Americans were recruited to participate on Mturk for \$.50.<sup>9</sup> Data from 13 cases were excluded from analyses *a priori* due to failing a key attention check (10 participants), and/or reporting difficulty viewing study materials (3 participants). The remaining 135 participants (63 women) ranged in age from 21 to 75 years ( $M=32.49$ ,  $SD=10.45$ ). It took participants on average 9.69 min to complete all survey materials. The average self-reported political orientation (1 = *very conservative*, 4 = *moderate*, 7 = *very liberal*) of our sample was skewed slightly liberal ( $M=5.21$ ,  $SD=1.41$ ). Participants were randomly assigned to either express outrage at third-party harm-doing (*third-party-directed outrage* condition) or not (*control* condition) in a between-subjects experimental design.

## Materials and procedure

First, participants read the same fabricated news article used in Study 1 which detailed the ubiquity of sweatshop labor conditions and their harmful effects in the developing world. Participants then rated the extent to which they believed "workers in developing countries were suffering as a result of sweatshop labor conditions" (1 = *not at all*, 7 = *very much*). A one-sampled *t* test revealed that participants responses to this item ( $M=6.49$ ,  $SD=.91$ ) were

<sup>9</sup> A sample size analysis using G\*Power (Faul et al. 2007) revealed that in order to ensure .80 power for our primary analysis, assuming a small to medium effect size we would need a total sample size of 124. Factoring in the exclusion rate from the previous studies we estimated that we would need to collect approximately 148 participants.



significantly higher than scale's midpoint (4),  $t(133)=31.71$ ,  $p<.001$ , indicating a general acknowledgment that sweatshop workers experience high levels of harm.

### Third-party harm-doing

Next, all participants completed a 4-item measure used in previous research (Rothschild et al. 2012, 2013) to assess the degree to which participants believed that *international corporations are guilty, to blame, at fault, and responsible* for the suffering of those working in sweatshops in the developing world. Responses were made on a 7-point scale (1=*not at all*, 7=*very much*) and were averaged to form composite third-party harm-doing scores ( $M_{grand} = 6.01$ ,  $SD = .97$ ;  $\alpha = .95$ ).

### Moral outrage manipulation

As in Study 3, participants were then randomly assigned to complete one of two surveys, comprising our primary experimental manipulation. Participants assigned to the *third-party-directed outrage* condition were presented with the same 6-item moral outrage measure used in Studies 2 and 3 ( $M_{grand} = 5.71$ ,  $SD = 1.32$ ;  $\alpha = .95$ ). Completing this measure gave participants the opportunity to express moral outrage at the third-party perpetrated harm.

In contrast, participants assigned to the *control* condition were not given the opportunity to express outrage at third-party harm-doing. Instead, participants completed a modified version of the Personal Need for Structure (PNS) scale (Thompson et al. 2001). Specifically, participants were asked to rate their level of agreement (1=*strongly disagree*, 7=*strongly agree*) with 6 statements related to the desire for simple structure (e.g., "I enjoy having a clear and structured mode of life";  $M = 3.77$ ,  $SD = .54$ ;  $\alpha = .69$ ).

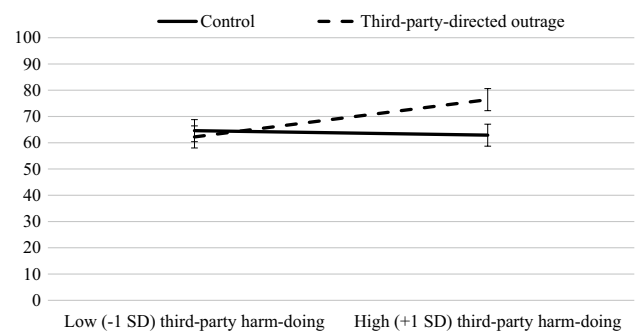
### Personal moral character measure

Finally, participants' completed the same measure of perceived personal morality used in Study 3 which asked participants to rank their own moral character relative to others they knew (0=*worse than all others*, 100=*better than all others*;  $M_{grand} = 66.57$ ,  $SD = 18.26$ ).

## Results

### Personal moral character

To test our primary hypothesis, we regressed personal moral character ratings onto the moral outrage manipulation (coded: *third-party-directed outrage* = 1, *control* = 0), perceived third-party harm-doing (continuous and centered),



**Fig. 8** Personal moral character as a function of perceived third-party harm-doing and opportunity to express third-party-directed outrage (Study 4). Black bars represent standard error

and the interaction of these predictors. As predicted, we observed a Moral outrage  $\times$  Third-party harm-doing interaction,  $\beta = .30$ ,  $b = 8.18$ ,  $SE = 3.11$ ,  $t(131) = 2.63$ ,  $p = .01$ , as well as marginal main effects of both predictors [*moral outrage*:  $\beta = .15$ ,  $b = 5.58$ ,  $SE = 3.09$ ,  $t = 1.81$ ,  $p = .07$ ; *Third-party harm-doing*:  $\beta = .16$ ,  $b = 3.07$ ,  $SE = 1.59$ ,  $t = 1.93$ ,  $p = .06$ ] (see Fig. 8 for the pattern of this interaction). The interaction remained significant when controlling for participants' comparative rankings of their standing on morally irrelevant positive traits,  $\beta = .25$ ,  $b = 6.84$ ,  $SE = 2.94$ ,  $t(130) = 2.33$ ,  $p = .02$ .

Consistent with predictions, among those who blamed international corporations for harmful sweatshop conditions (+1 *SD*), the opportunity to express moral outrage at corporate harm-doing (vs. not) led to significantly higher personal moral character ratings,  $\beta = .37$ ,  $b = 13.53$ ,  $SE = 4.28$ ,  $t = 3.16$ ,  $p = .002$ . In contrast, among participants who perceived international corporations as less culpable for sweatshop labor conditions (-1 *SD*), the opportunity to express third-party-directed outrage had no significant effect on personal moral character ratings,  $\beta = -.07$ ,  $b = -2.39$ ,  $SE = 4.28$ ,  $t = -.56$ ,  $p = .58$ .

Furthermore, simple slopes analysis revealed that in the *third-party-directed outrage* condition, the perceived harm done by international corporations was positively and significantly related to personal moral character ratings,  $\beta = .39$ ,  $b = 7.29$ ,  $SE = 2.24$ ,  $t = 3.26$ ,  $p = .001$ . In contrast, when individuals were not given the opportunity to express outrage, perceived third-party harm-doing was unrelated to personal moral character ratings,  $\beta = -.05$ ,  $b = -.88$ ,  $SE = 2.16$ ,  $t = -.41$ ,  $p = .68$ .<sup>10</sup>

<sup>10</sup> As in Study 3 we conducted correlational analyses to test for the hypothesized relationship between moral outrage at third-party harm-doing and perceived personal moral character. Consistent with the previous study, this analysis revealed a significant positive correlation between moral outrage scores and personal moral character ratings following a moral identity threat,  $r(67) = .25$ ,  $p = .05$ .

## Discussion

Supporting the primary prediction, among those who perceived high levels of corporate harm-doing, it was the opportunity to express moral outrage at corporate harm-doers that bolstered participant's perceived personal morality. Furthermore, this willingness to blame corporations predicted more favorable perceptions of one's own moral character only when participants were able to express outrage. This suggests that the observed effects are not merely the consequence of comparing one's own harm-doing against more egregious harm perpetrated by a third-party. Rather, these findings support our claim that directing moral outrage against a third-party can serve to defend one's moral identity.

To the extent that some expressions of outrage seem to provide moral identity maintenance, we would expect these expressions of outrage to be attenuated if one's own moral status is affirmed. Study 5 was designed to test this prediction.

## Study 5

Study 5 had three main goals. First, because past research has focused primarily on the role of collective guilt in moral outrage (e.g., Rothschild et al. 2013), we sought to test whether affirmations of *personal* moral identity could attenuate the effects of *collective* guilt on moral outrage. We did this by first assessing (American) participants' guilt for the harm that American consumers caused to sweatshop workers. We then assessed outrage at corporations for harmful sweatshop practices. In line with our previous findings, we hypothesized that when no affirmation was available:

**H6a:** *Greater collective guilt would predict increased feelings of moral outrage at corporate harm-doing.*

Previous research has shown that people can compensate for a threat to their moral identity by affirming their moral status, even in an unrelated context (Rothschild et al. 2012). A second goal of Study 5 was to test whether expressions of moral outrage could similarly be moderated by attempts to bolster participants' perceived moral identity. Based on the idea that defensive outrage is ultimately grounded in the desire to restore a perceived moral identity, we hypothesized that:

**H6b:** *A personal moral identity boost would attenuate guilt-induced expressions of outrage, effectively eliminating the relationship between guilt over ingroup harm-doing and outrage at third-party perpetrated harm.*

We tested this by manipulating whether participants were given the opportunity to affirm their personal moral

identity (*moral affirmation* condition) or not (*control affirmation* condition) after assessing guilt over ingroup harm-doing, but prior to assessing outrage at corporate harm-doing. If expressions of outrage are stifled by a moral affirmation, this would provide strong evidence that defensive outrage is grounded in concerns about one's own moral identity, rather than maintaining justice.

Some have criticized previous research for failing to differentiate between moral outrage and empathic anger (Batson 2011), broadly defined as anger at undeserved harm against a cared-for other (Hoffman 2000). Study 5 also addressed the possibility that differences in outrage at corporate harm-doing may reflect differences in participant's tendency to empathize with the perceived victims of this harm. We did this by measuring and statistically controlling for the extent to which participants empathized with those working in sweatshop labor conditions in the developing world.

## Method

Ninety-nine American adults were recruited to participate on Mturk for \$.75.<sup>11</sup> Data from 18 cases was excluded from analyses *a priori* due to failing an attention check (16 participants) and/or reported difficulty viewing study materials (2 participants). The remaining 81 participants (35 women) ranged in age from 18 to 68 years ( $M=31.75$ ,  $SD=10.55$ ). It took participants on average 12.82 min to complete all survey materials. As in the previous studies, the sample was somewhat liberal ( $M=4.56$ ,  $SD=1.64$ ). Participants were randomly assigned to either the moral identity or control affirmation and materials were presented in the order below.

## Materials and procedure

Participants first read the same fabricated news article on labor exploitation used in Study 1, which highlighted the horrific conditions faced by sweatshop workers in the developing world. Participants then rated the extent to which they believed "workers in developing countries were suffering as a result of sweatshop labor conditions" (1 = *not at all*, 7 = *very much*). A one-sample *t*-test revealed that participants' responses to this item ( $M=6.43$ ,  $SD=.81$ ) were significantly higher than scale's midpoint (4),  $t(80)=27.12$ ,  $p<.001$ , indicating a general acknowledgment that sweatshop workers experience high levels of harm.

<sup>11</sup> A sample size analysis using G\*Power (Faul et al. 2009) revealed that in order to ensure .80 power for our primary analysis, assuming a small to medium effect size we would need a total sample size of 89. Factoring in the exclusion rate from the previous studies we estimated that we would need to collect approximately 100 participants.

### Collective guilt

Participants then completed a 3-item measure of collective guilt used in previous research (Doosje et al. 1998; Rothschild et al. 2013) to assess feelings of guilt for the harm caused by one's own group. Specifically, participants indicated the extent to which, as Americans, they felt *guilty*, *regretful*, and *apologetic* for harm American Consumers have caused workers in the developing world. Responses were made on a 7-point scale (1 = *not at all*, 7 = *very much*) and were averaged to form composite collective guilt scores ( $M_{\text{grand}} = 4.25$ ,  $SD = 1.71$ ;  $\alpha = .92$ ).

### Exposure to third-party harm-doer

We then presented all participants with a third-party harm-doer by having participants read an ostensible news article titled "Apple's Factories Still 'Sweatshops' says Watchdog Group." The article described exploitative and "inhumane" labor practices uncovered at Apple Inc.'s Chinese factories which included "denying workers' basic human needs, such as allowing bathroom breaks, sufficient rest, and access to proper nutrition". The article unambiguously blamed Apple Inc., stating "despite being aware of a multitude of labor abuses in these factories, Apple Inc. failed to take action to stop these violations. As such they bear the responsibility for the suffering of thousands of workers."

### Affirmation manipulation

Next, participants completed a writing task that was purported to be a personality assessment. This task, taken from previous research (Rothschild et al. 2012) constituted our affirmation manipulation. Participants assigned to the *moral identity affirmation* condition responded to the following writing prompt: "In a few sentences briefly describe something about yourself that makes you feel like a good and decent person." Participants in the *control affirmation* condition responded to the following writing prompt: "In a few sentences please briefly describe something in your life that you have complete control over." An inspection of participants' written responses revealed that all participants wrote at least one sentence and no participants wrote about issues directly pertaining to sweatshop labor.

### Empathy measure

All participants completed a modified version of a 4-item empathy scale used in previous research (Pagano and Huo 2007; Rothschild et al. 2013) to assess the extent to which participants felt *empathy*, *sympathy*, and *compassion* for sweatshop workers and were *moved* by their plight. Responses were made on a 7-point scale (1 = *not at all*,

7 = *very much*) and were averaged to form composite empathy scores ( $M_{\text{grand}} = 5.38$ ,  $SD = 1.52$ ;  $\alpha = .97$ ).

### Moral outrage

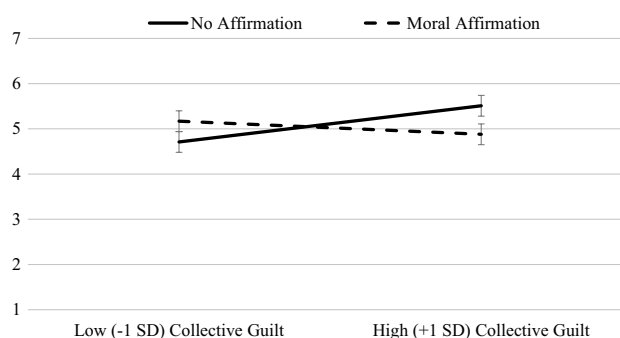
Participants completed the same moral outrage measure used in Studies 1 and 4 to assess the extent to which participants feel anger towards international corporations for harm done to sweatshop workers in the developing world ( $M_{\text{grand}} = 5.14$ ,  $SD = 1.43$ ;  $\alpha = .97$ ).

## Results

### Moral outrage

To test our primary hypothesis, we regressed moral outrage onto affirmation manipulation (coded: *moral identity* = 1, *control* = 0), collective guilt (continuous and centered) and their interaction. We included empathy as additional continuous predictor. Importantly, as predicted this analysis yielded a significant Affirmation  $\times$  Collective Guilt interaction,  $\beta = -.25$ ,  $b = -.31$ ,  $SE = .10$ ,  $t(76) = -3.22$ ,  $p = .002$ , but no significant effects for the individual predictors [*collective guilt*:  $t = 1.09$ ,  $p = .28$ ; *affirmation*:  $t = -.38$ ,  $p = .70$ ] (see Fig. 9 for the pattern of this interaction). The Affirmation  $\times$  Collective Guilt interaction remained significant when empathy was not included in the model,  $\beta = -.81$ ,  $b = -.53$ ,  $SE = .16$ ,  $t(77) = -3.35$ ,  $p = .001$ .

Consistent with predictions a simple slopes analysis revealed that in the *control affirmation* condition, collective guilt was positively related to moral outrage at corporate harm-doing,  $\beta = .28$ ,  $b = .23$ ,  $SE = .08$ ,  $t = 3.03$ ,  $p = .003$ . In contrast, in the *moral affirmation* condition, collective guilt was unrelated to moral outrage,  $\beta = -.10$ ,  $b = -.08$ ,  $SE = .07$ ,  $t = -1.13$ ,  $p = .26$ . Also supporting predictions, comparison of the high-collective guilt participants (+1 SD) demonstrated that those given the opportunity to affirm



**Fig. 9** Moral outrage as a function of collective guilt and affirmation condition (Study 5). Black bars represent standard error

their moral identity (vs. control) showed reduced moral outrage  $\beta = -.21$ ,  $b = -.62$ ,  $SE = .24$ ,  $t = -2.62$ ,  $p = .01$ . In contrast, among participants low in collective guilt ( $-1$  SD) those who affirmed their moral identity (vs. control) reported marginally more moral outrage compared to those in the control affirmation condition,  $\beta = .16$ ,  $b = -.45$ ,  $SE = .23$ ,  $t = 1.96$ ,  $p = .05$ .

## Discussion

Supporting predictions, a moral affirmation attenuated moral outrage among those who reported high levels of collective guilt, and effectively eliminated the relationship between guilt and outrage. Furthermore, we found this pattern despite the fact that the moral affirmation in this study was completely irrelevant to the source of collective guilt. Finally, participants' tendency to empathize with the victims of sweatshop labor practices did not account for the observed data. This helps rule out the possibility that these effects represent differences in empathic anger. In combination with our prior studies, these results confirm that a direct moral affirmation substantially reduces expressions of outrage, further supporting the proposed defensive account.

One unexpected finding was that among those who reported low levels of collective guilt, those who affirmed their moral identity expressed marginally more outrage at corporate harm-doing than those in the control affirmation condition. It is possible that those low in guilt may have experienced the affirmation of their moral identity as a means of establishing moral credentials. Previous research has shown that establishing moral credentials allows a person to engage in unethical behavior without experiencing distress (Self-licensing effect; see Merritt et al. 2010). This raises the possibility that among participants who did not perceive their group's contribution to sweatshop labor conditions as a moral identity threat, a moral affirmation may have provided the moral credentials necessary to license the free expression of outgroup hostility.

## General discussion

Five studies supported the proposed guilt-driven account of defensive moral outrage whereby self-focused guilt over personal or collective harm-doing motivates expressions of moral outrage at a third-party. Study 1 showed that guilt in response to salient personal harm-doing predicted an increased desire to punish a third-party through increased moral outrage at that target. Building on Study 1, Study 2 manipulated responsibility for harm-doing to show that salient ingroup (vs. outgroup) harm-doing elicited outrage

at a third-party through increased feelings of personal guilt. Studies 2 and 3 directly tested the defensive function of moral outrage, finding that the opportunity to express third-party-directed outrage effectively attenuated guilt (Study 2) and bolstered personal moral identity (Study 3) following the threat of ingroup immorality. Study 4 showed that exposure to a third-party's immoral action was insufficient to bolster evaluations of one's own moral character in the absence of an opportunity to express moral outrage. Finally, Study 5 showed that guilt-induced moral outrage was attenuated when participants were provided with an alternative means of bolstering their moral identity, even in an unrelated context.

We took steps to rule out a number of potential alternative explanations for these results. Most notably, Study 4 was designed to disentangle the effects of expressing third-party-directed moral outrage from the ability to make a downward social comparison between oneself and a comparably more egregious third-party. We also attempted to test alternative explanations by assessing and statistically controlling for an array of extraneous variables. For example, Study 1 showed that the association between guilt for personal harm-doing and outrage at third-party harm-doing was not reducible to variability in participants' general negative affect. Similarly, we took steps to test the possibility that our obtained effects may have reflected variations in common "nonmoral" forms of anger (i.e., personal anger and empathic anger; Batson 2011) rather than moral outrage. In Study 2 we attempted to exclude the potential influence of personal anger by assessing and controlling for the extent to which participants believed that they were, or might become, victims of climate change. In Study 5, we attempted to exclude the influence of empathic anger by assessing and controlling for participants' tendency to empathize with the workers undeservedly harmed by corporations' sweatshop labor practices. Even though the contexts in these studies are highly charged political topics, we found that controlling for ideology did not explain the observed results throughout.

These findings are consistent with recent research showing that outgroup-directed moral outrage can be elicited in response to perceived threats to the ingroup's moral status (Rothschild et al. 2013; Täuber and van Zomeren 2013). However, by assessing mediating and moderating variables along with downstream consequences the present studies go beyond previous literature. We found consistent evidence that outrage driven by moral identity concerns serves to compensate for the threat of personal or collective immorality. This research also contributes more generally to the moral emotions literature by illuminating a link between guilt and self-serving expressions of outrage that reflect a kind of "moral hypocrisy" (Batson et al. 1999), or at least a nonmoral form of anger with a moral façade. In



particular, the attenuation of elevated outrage in Studies 5 highlights the disingenuous nature of these expressions and exposes one's own moral identity, rather than justice, as the central motivating concern in these cases.

However, it should be clear that we do not propose that *all* expressions of moral outrage are the result of deep-seated guilt or defensive attempts to maintain a moral image. Surely there exist forms of moral outrage genuinely motivated by a desire to promote social justice or to combat oppression. Our central claim is simply that outrage *can* be defensive and *can* be motivated by underlying feelings of guilt in order to bolster a moral self-concept *when* people's moral identity is threatened. Distinguishing the motives behind moral outrage presents opportunities for research exploring the forms and functions of this phenomenon. For example, we might expect that more genuine (i.e., justice-oriented) forms of moral outrage are likely to motivate more high-investment behaviors over a longer period of time to combat injustice. In contrast, defensive outrage is likely to elicit only as much action as is necessary to make the outraged party feel moral once more. In the context of online activism, one wonders how much short-lived movements like “#StopKONY” (Bellantoni and Polantz 2013), “#GamerGate” (Hathaway 2014) and others communities owe their existence and perpetuation to moral identity maintenance.

### Limitations and future directions

While our results consistently support the proposed defensive function of moral outrage, they do not speak to *how* outrage serves this function. However, past research offers some possibilities. For example, recent research suggests an unaffected observer's tendency to harshly punish a third-party may serve as a signal of one's own selflessness, honesty, and trustworthiness to others (Jordan et al. 2016). In support of this claim, Jordan and colleagues found that punishing third-party perpetrators was viewed by others as a sign of trustworthiness and the desire to punish was attenuated when participants were provided with other means of communicating their moral worth. Other research has suggested that harsh views of moral transgressors may be used to distance oneself from ethical misconduct (Barkan et al. 2012). This suggests that participants in the present studies may have used expressions of third-party-directed outrage as a means of signaling a moral identity to others or distancing themselves from salient personal or collective transgressions that might otherwise threaten their moral reputation. However, such explanations, grounded in the communication and management of one's public self-image, fail to offer a compelling reason why private feelings of guilt should predict moral outrage and more importantly, why the expression of such outrage should alleviate

felt guilt, even in the context of an anonymous online study. As such, future research is needed not only to confirm the replicability of the observed process and identify its potential boundary conditions, but also to illuminate how it operates.

The current studies provide strong evidence that subjective feelings of guilt can spur expressions of third-party-directed outrage to restore a moral identity. However, the current research paradigm is limited insofar as it focuses on emotional responses to past or ongoing acts of harm-doing. Research finds that in some cases the mere anticipation that one *might* experience guilt over future harm-doing can motivate guilt avoidance strategies, such as the intention to perform prosocial actions or avoid bad behavior (e.g., Sandberg and Connor 2008; Wang and McClung 2012). This raises the possibility that the mere *anticipation* of guilt over one's potential future harm-doing may be sufficient to elicit third-party-directed outrage as a preemptive guilt evasion strategy. If true, this would suggest that whereas subjective feelings of guilt can drive defensive displays of moral outrage (as shown in the present research), the actual experience of guilt in the present may not be necessary. Future research is needed to further clarify the precise role played by guilt (experienced and anticipated) in the production of defensive outrage.

Importantly, the current studies' emphasis on moral identity concerns does not preclude the potential that guilt may serve other functional capacities. For example, research suggests that guilt can play an important function in facilitating and maintain social relationships (e.g., Baumeister et al. 1994). This work views guilt as an evolved mechanism which signals that one has caused harm or distress to a relationship partner and motivates efforts to repair the damaged relationship (Keltner et al. 2006). Combining this insight with the current project raises the possibility that guilt-induced expressions of moral outrage may also serve a relationship maintenance function: Expressions of outrage may serve to mend a relationship by signaling one's concern for, and commitment to fighting on behalf of an injured party. Other researchers have theorized that moral outrage at a third-party may foster a sense of social cohesion and promote a shared identity between advantaged and disadvantaged groups against that transgressor (Thomas et al. 2009). Future research is needed to explore conditions in which moral outrage may be employed for social benefit.

The scope of the present research is also limited by its exclusive focus on the relationship between guilt and outrage. One important direction for future research will be to shed light on the potential role of other emotions in motivating moral outrage. Shame for example is another core “moral emotion” which is similar to guilt insofar as it is caused by personal or collective immorality (Tangney et al.



2007). However, unlike guilt, shame is purely self-evaluative, rather than reflecting an evaluation of some particular moral wrong (e.g., feeling that one is worthless rather than that one did something improper; Tangney 1992). Would shame contribute to moral outrage in the same way as guilt? At first blush this seems unlikely: Because guilt evaluates an action as wrong, it is desirable to hold others accountable for this action. In contrast, shame about one's own moral status is not easily attributed to international corporations, immigrants, or other outside sources. However, recent research has questioned the conceptual boundaries between guilt and shame (Gausel and Brown 2012), and shown that shame can predict self-defensive motivations (Gausel et al. 2012), particularly when people perceive a risk to their social-image and fear condemnation from others (Gausel et al. 2016). This suggests that shame might be more likely to predict defensive outrage when social image concerns are salient.

We also provided a fixed target (e.g., corporations) for participants' outrage in the present studies in order to assess the specific consequences of moral outrage while reducing "real-world" noise by isolating the phenomenon. However, it seems reasonable that people might actively seek or spontaneously generate targets of outrage as desired. Factors influencing the spontaneous selection of targets of outrage may include salience (e.g., from media exposure), prior exposure (as frequent targets of outrage), and the motivation to maintain a consistent worldview. Glick (2005) argues that people turn to shared ideologies when motivated to identify outgroups that can be blamed and punished for negative outcomes. Furthermore, recent research suggests that the inflated perceptions of third-party harm-doing can be motivated by the need to justify feelings of outrage (Thomas et al. 2016). This raises the possibility that people may be able to find and/or "create" outrage-worthy targets when motivated to defend their moral status.

Additionally, while we believe that our decision to investigate the complex real world issues of consumer and corporate harm-doing is a strength of the present research, it does make it difficult for us to definitively rule out the possibility that some participants might have viewed themselves as contributing to corporate harm-doing. Insofar as participants saw themselves as complicit in a larger system of harm-doing which encompasses both the self and the corporations, participants' outrage at corporate harm-doing may have reflected ingroup-directed or self-focused anger, which has been shown to have a strong to moderate association with guilt (e.g., Iyer et al. 2007). For instance, if participants in Study 5 felt implicated in Apple Inc.'s sweatshop labor practices through their use of Apple products, their expressions of outrage at Apple Inc. might have reflected some degree of self-focused anger as opposed to outgroup-directed outrage. However, if participants' were

exhibiting self or ingroup-focused anger in the present studies it is unclear why these expressions would reduce subsequent feelings of guilt as was observed in Study 2.

Ultimately, while our context has practical implications, it does so at some cost in its ability to tease apart the roles of self vs. corporate responsibility. Regardless of whether participants' outrage at corporate harm-doing was self- or other-focused in the present studies, our findings offer compelling evidence of a defensive process and a novel account of an emotion previously assumed to reflect a motive for justice restoration (Harth et al. 2011, 2013; Thomas et al. 2009). However, future research might overcome this limitation by focusing on a third-party whose harm-doing is more clearly disconnected from participants' own behavior, or by directly assessing the perceived overlap between the participants and a third-party harm-doer.

Finally, on a more practical note, our studies were limited by an online MTurk sampling method that yielded a number of low-quality respondents, marked by inattention and evidence of low conscientiousness. Although series of a priori plans were implemented to account for, identify, and exclude these participants (e.g., use of attention check items, timed responses, probing questions), we cannot be certain that all such cases were excluded or that these exclusions might have narrowed our sample along other dimensions. As such, we advise that extensions of this work employ alternative methods of sampling in order to draw firmer conclusions about the strength and generalizability of the obtained effects.

## Implications

An important implication of this research is the possibility that at least some manifestations of moral outrage may be driven by feelings of guilt. In a political climate that has tended toward greater political polarization in recent years (PEW 2014), it is increasingly important to find ways of defusing virulent or distracting expressions of outrage. An acknowledgement of the defensive motives that underlie certain forms of outrage points toward at least one source of strain on clear-headed political discourse. Importantly, this research also gestures toward a corrective: To the extent that individuals and groups can find positive ways of maintaining a moral identity, they may be able to diffuse defensive moral outrage.

It should be clear, however, that moral outrage is not inherently problematic nor to be avoided. Collective outrage can be a powerful motivating force for social change. Furthermore, even if feelings of outrage are grounded in underlying guilt, it does not necessarily preclude expressions of such outrage from motivating prosocial or even self-sacrificial actions. For instance, defensive outrage at corporate malfeasance might lead to support for a corporate

boycott in which one sacrifices one's ability to obtain desired goods in order to pressure corporate reform. However, to the extent that moral outrage aims not at positive social change, but merely at personal moral affirmation, one may at least question its value for meaningful social reform.

**Funding** These studies were funded by the authors.

#### Compliance with ethical standards

**Conflict of interest** Zachary Rothschild declares that he has no conflict of interest. Lucas Keefer declares that he has no conflict of interest.

**Ethical approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

## References

- Barkan, R., Ayal, S., Gino, F., & Ariely, D. (2012). The pot calling the kettle black: Distancing response to ethical dissonance. *Journal of Experimental Psychology, 141*, 757–773.
- Batson, C. D. (2011). What's wrong with morality? *Emotion Review, 3*, 230–236.
- Batson, C. D., Thompson, E. R., Seufferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology, 77*, 525–537.
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1994). Guilt: An interpersonal approach. *Psychological Bulletin, 115*, 243–267.
- Bellantoni, C., & Polantz, K. (2013, April 19). One year later: What happened to #stopkony? Retrieved from <http://www.pbs.org/newshour/rundown/one-year-later-what-happened-to-stopkony/>
- Branscombe, N. R., Doosje, B., & McGarty, C. (2002). Antecedents and consequences of collective guilt. In D. M. Mackie & E. R. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 49–66). New York: Psychology Press.
- Branscombe, N. R., & Miron, A. M. (2004). Interpreting the ingroup's negative actions toward another group: Emotional reactions to appraised harm. In L. Z. Tiedens & C. W. Leach (Eds.), *The social life of emotions* (pp. 314–335). New York: Cambridge University Press.
- Branscombe, N. R., Slugoski, B., & Kappen, D. M. (2004). Collective guilt: what it is and what it is not. In N. R. Branscombe & B. Doosje (Eds.), *Collective guilt: International perspectives* (pp. 16–34). Cambridge: Cambridge University Press.
- Doosje, B., Branscombe, N. R., Spears, R., & Manstead, A. S. R. (1998). Guilty by association: When one's group has a negative history. *Journal of Personality and Social Psychology, 75*, 872–886.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
- Fritz, M. S., & McKinnon, D. P. (2007). Required samples size to detect the mediated effect. *Psychological Science, 18*, 233–239.
- Gausel, N., & Brown, R. (2012). Shame and guilt—Do they really differ in their focus of evaluation? Wanting to change the self and behavior in response to ingroup immorality. *The Journal of Social Psychology, 152*, 547–567.
- Gausel, N., Leach, C. W., Vignoles, V. L., & Brown, R. (2012). Defend or repair? Explaining responses to in-group moral failure by disentangling feelings of shame, rejection, and inferiority. *Journal of Personality and Social Psychology, 102*, 941.
- Gausel, N., Vignoles, V. L., & Leach, C. W. (2016). Resolving the paradox of shame: Differentiating among specific appraisal-feeling combinations explains pro-social and self-defensive motivation. *Motivation and Emotion, 40*, 118–139.
- Glick, P. (2005). Choice of scapegoats. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 244–261). Malden, MA: Blackwell Publishing.
- Gupta, P. (2012, January 27). Tim Cook responds to Foxconn worker abuse report: Apple not turning 'blind eye'. Retrieved from [http://www.huffingtonpost.com/2012/01/27/tim-cook-foxconn\\_n\\_1237341.html](http://www.huffingtonpost.com/2012/01/27/tim-cook-foxconn_n_1237341.html)
- Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 852–870). Oxford: Oxford University Press.
- Harth, N. S., Hornsey, M., & Barlow, F. K. (2011). Emotional responses to rejection of gestures of intergroup reconciliation. *Personality and Social Psychology Bulletin, 37*, 815–829.
- Harth, N., Leach, S. C. W., & Kessler, T. (2013). Guilt, anger, and pride about ingroup environmental behavior: Different emotions predict distinct intentions. *Journal of Environmental Psychology, 34*, 18–26.
- Harvey, R. D., & Oswald, D. L. (2000). Collective guilt and shame as motivation for white support of black programs. *Journal of Applied Social Psychology, 30*, 1790–1811.
- Hathaway, J. (2014, October 10). What Is Gamergate, and Why? An explainer for Non-Geeks. Retrieved from <http://gawker.com/what-is-gamergate-and-why-an-explainer-for-non-geeks-1642909080>
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. New York: Cambridge University Press.
- Iyer, A., Leach, C. W., & Crosby, F. J. (2003). White guilt and racial compensation: The benefits and limits of self-focus. *Personality and Social Psychology Bulletin, 29*, 117–129.
- Iyer, A., Schmader, T., & Lickel, B. (2007). Why individuals protest the perceived transgressions of their country: The role of anger, shame, and guilt. *Personality and Social Psychology Bulletin, 33*, 572–587.
- Jordan, J., Hoffman, M., Bloom, P., & Rand, D. G. (2016). Third-party punishment as costly signal of trustworthiness. *Nature, 530*, 473–476.
- Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin, 37*, 701–713.
- Keltner, D., Haidt, J., & Shiota, M. N. (2006). Social functionalism and the evolution of emotions. In M. Schaller, J. A. Simpson & D. T. Kendrick (Eds.), *Evolution and Social Psychology* (pp. 115–142). New York: Psychology Press.

- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of ingroups. *Journal of Personality and Social Psychology, 93*, 234–249.
- Leach, C. W., Snider, N., & Iyer, A. (2002). Poisoning the consciences of the fortunate: The experience of relative advantage and support for social equality. In I. Walker & H. J. Smith (Eds.), *Relative deprivation* (pp. 136–163). New York: Cambridge University Press.
- Lodewijkz, H. F. M., Kersten, G. L. E., & van Zomeren, M. (2008). Dual pathways to engage in “silent marches” against violence: Moral outrage, moral cleansing and modes of identification. *Journal of Community and Applied Psychology, 18*, 153–167.
- Luttrel, A., Petty, R. E., Briñol, P., & Wagner, B. C. (2016). Making it moral: Merely labeling an attitude as moral increases its strength. *Journal of Experimental Social Psychology, 65*, 82–93.
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*, 633–644.
- McGarty, C., Pedersen, A., Leach, C. W., Mansell, T., Waller, J., & Bliuc, A. (2005). Group-based guilt as a predictor of commitment to apology. *British Journal of Social Psychology, 44*, 659–680.
- Merritt, A. C., Effron, D., & Monin, B. (2010). Moral licensing: When being good frees us to be bad. *Social and Personality Psychology Compass, 4*, 344–357.
- Miron, A. M., Branscombe, N. R., & Biernat, M. R. (2010). Motivated shifting of justice standards. *Personality and Social Psychology Bulletin, 36*, 768–779.
- Monin, B., & Jordan, A. H. (2009). Dynamic moral identity: A social psychological perspective. In D. Narvaez & D. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 341–354). Cambridge: University Press.
- Montada, L., & Schneider, M. (1989). Justice and emotional reactions to the disadvantaged. *Social Justice Research, 3*, 313–344.
- Nisan, M. (1991). The moral balance model: Theory and research extending our understanding of choice and deviation. In W. M. Kurtines & J. L. Gerwitz (Eds.), *Handbook of moral behavior and development* (Vol. 3, pp. 213–249). Hillsdale, NJ: Lawrence Erlbaum.
- O'Connor, L. E. (2010). Guilt. In I. B. Weiner & W. E. Craighead (Eds.), *Corsini encyclopedia of psychology*. Hoboken, NJ: Wiley.
- Pagano, S., & Huo, Y. J. (2007). The role of moral emotions in predicting support for political actions in post-war Iraq. *Political Psychology, 28*, 227–255.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879–891.
- PEW, (2014, June 12). Political polarization in the American republic. Retrieved from <http://www.people-press.org/2014/06/12/political-polarization-in-the-american-public/>
- Rothschild, Z. K., Landau, M. J., Keefer, L. A., & Sullivan, D. (2015). Another's punishment cleanses the self: Evidence for a moral cleansing function of punishing transgressors. *Motivation and Emotion*. doi:10.1007/s11031-015-9487-9. (Advanced online publication)
- Rothschild, Z. K., Landau, M. J., Molina, L. E., Branscombe, N. R., & Sullivan, D. (2013). Displacing blame over the ingroup's harming of a disadvantaged group can fuel moral outrage at a third-party scapegoat. *Journal of Experimental Social Psychology, 49*, 898–906.
- Rothschild, Z. K., Landau, M. J., Sullivan, D., & Keefer, L. A. (2012). A dual-motive model of scapegoating: Displacing blame to reduce guilt or increase control. *Journal of Personality and Social Psychology, 102*, 1148–1163.
- Sandberg, T., & Conner, M. (2008). Anticipated regret as an additional predictor in the theory of planned behaviour: A meta-analysis. *British Journal of Social Psychology, 47*, 589–606.
- Stewart, T. L., Latu, I. M., Branscombe, N. R., & Denney, H. T. (2010). Yes we can! Prejudice reduction through seeing (inequality) and believing (in social change). *Psychological Science, 21*, 1557–1562.\*\*
- Tangney, J. P. (1992). Self-conscious emotions: The self as a moral guide. In T. Abraham, D. A. Staple & J. V. Wood (Eds.), *Self and motivation: Emerging psychological perspectives* (pp. 97–117). Washington, DC: American Psychological Association.
- Tangney, J. P. (1995). Recent advances in the empirical study of shame and guilt. *The American Behavioral Scientist, 38*, 1132–1145.
- Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology, 58*, 345–372.
- Tarrant, M., Branscombe, N. R., Warner, R., & Weston, D. (2012). Social identity and perceptions of torture: It's moral when we do it! *Journal of Experimental Social Psychology, 48*, 513–518.
- Täuber, S., & van Zomeren, M. (2013). Outrage towards whom? Threats to moral group status impede striving to improve via outgroup-directed outrage. *European Journal of Social Psychology, 43*, 149–159.
- Thomas, A. J., Stanford, P. K., & Sarnecka, B.W. (2016). No child left alone: Moral judgments about parents affect estimates of risk to children. *Collabra, 2*(1), 10. DOI: [10.1525/collabra.33](https://doi.org/10.1525/collabra.33).
- Thomas, E. F. (2005). The role of social identity in creating positive beliefs and emotions to motivate volunteerism. *Australian Journal on Volunteering, 10*, 45–52.
- Thomas, E. F., McGarty, C., & Mavor, K. I. (2009). Transforming “Apathy into movement”: The role of prosocial emotions in motivating action for social change. *Personality and Social Psychology Review, 13*, 310–333.
- Thompson, M. M., Naccarato, M. E., Parker, K. C. H., & Moskowitz, G. (2001). The Personal Need for Structure (PNS) and Personal Fear of Invalidity (PFI) scales: Historical perspectives, present applications and future directions. In G. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 19–39). Mahwah, NJ: Erlbaum.
- Van Bavel, J. J., Packer, D. J., Haas, I. J., & Cunningham, W. A. (2012). The importance of moral construal: Moral versus non-moral construal elicits faster, more extreme, universal evaluations of the same actions. *PLoS ONE, 7*(11), e48693.
- Vidmar, N. (2000). Retribution and revenge. In J. Sanders & V. L. Hamilton (Eds.), *Handbook of justice research in law* (pp. 31–63). New York: Kluwer.
- Wang, X., & McClung, S. R. (2012). The immorality of illegal downloading: The role of anticipated guilt and general emotions. *Computers and Human Behavior, 28*, 153–159.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90*, 245–271.