

Word count: 9998

The Benefits of Collective Responsibility: How Ingroup Reputation Concern Motivates
Prosociality in Intergroup Contexts

Peter Kardos¹

Bernhard Leidner²

Emanuele Castano³

Brian Lickel²

1: Bloomfield College

2: University of Massachusetts Amherst

3: New School for Social Research

Correspondence:

Peter Kardos

Division of Social & Behavioral Science

Bloomfield College

59 Fremont Street, Room 205, Bloomfield, NJ 07003

peter_kardos@bloomfield.edu

COLLECTIVE RESPONSIBILITY

Abstract

Collective responsibility processes have been investigated from the perspectives of the outgroup (e.g. collective blame) and the ingroup (e.g. collective guilt). This paper extends theory and research on collective responsibility with a third perspective, namely that of the individual actor whose behavior triggers the attribution of collective blame. Four experiments tested the hypotheses that collective responsibility processes influence the individual actors' appraisals, emotions and behavior. The possibility of collective blame for their individual action prompted more prosocial behavior among participants (Experiment 1). Participants also experienced more ingroup reputation concern and in turn more negative emotions (Experiment 2-4) for a past wrongdoing if it could reflect negatively on the ingroup in the eyes of outgroups. The increased negative emotions then motivated participants to improve the ingroup's image (Experiment 4). The effects were further moderated by perceived ingroup entitativity (Experiment 3).

Keywords: collective responsibility; collective blame; intergroup context; entitativity; group reputation

COLLECTIVE RESPONSIBILITY

The Benefits of Collective Responsibility: How Ingroup Reputation Concern Motivates Prosociality in Intergroup Contexts

Individual wrongdoing often has group level implications. When a wrongdoer's membership in a social group is salient, outgroup members may not only judge the individual actor for his or her actions, but also blame or even retaliate against the individual's entire group (Lickel, Schmader, & Hamilton, 2003). While members of the wrongdoer's group sometimes support the wrongdoer's actions, they often react with feelings of guilt, shame, and anger (Lickel, Schmader, & Spanovic, 2007), and may try to protect the ingroup's reputation by sanctioning or excluding the wrongdoer from the group (Marques, Yzerbyt, & Leyens, 1988). To date, the mechanisms underlying collective responsibility have been investigated from these outgroup and ingroup perspectives, but not from the perspective of the individual actor whose behavior is the initial trigger of the collective responsibility process.

We extend theory and research on collective responsibility by investigating this third perspective of the collective responsibility process. We argue that people are generally aware that in certain contexts their individual actions may lead to collective blame. We propose that in contexts in which group memberships are salient and the ingroup might therefore be blamed for an individual actor's wrongdoing by an outgroup, the individual actor will likely appraise their past or anticipated behavior in terms of its potential effect on the ingroup's reputation. This appraisal should then motivate the individual to engage in more positive, prosocial behavior, either simply to avoid

COLLECTIVE RESPONSIBILITY

damaging the ingroup's reputation or to repair damage that possibly has already been caused by past behavior.

Past research on collective blame: Outgroup and ingroup perspectives

People tend to perceive themselves and others in terms of group membership (Fiske & Taylor, 2013). This tendency is strongest in intergroup contexts, which highlight social categories and in which people perceive themselves and others as members of groups instead of unique individuals (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). In these contexts, an individual's behavior may evoke attributions of collective blame from both outgroup and ingroup members.

Outgroup perspective. Research on social cognition and social identity demonstrates that people often generalize from an actor's behavior and characteristics to the actor's entire group (Crawford, Sherman, & Hamilton, 2002; Henderson-King & Nisbett, 1996; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; Brambrilla, Hewstone, & Colucci, 2013). For example, Crawford et al., (2002) showed that outsiders make trait generalizations from the individual group member's actions to the other group members. Such person to group generalization can be understood as an instance of collective responsibility because the outgroup denounces the entire group in order to explain an individual member's behavior Lickel and Onuki (2015) proposed that. In addition to these dispositional inferences, outside perceivers also often draw causal connections between an individual group member's actions and others in the group. Research (e.g. Denson, Lickel, Curtis, Stenstrom, & Ames, 2006; Lickel et al., 2003) shows that outsiders often believe that other group members indirectly encouraged or benefited from the action of the individual group member or that those other group members failed in a

COLLECTIVE RESPONSIBILITY

responsibility to monitor and control the blameworthy group member's behavior. Thus, when an individual group member commits a blameworthy act, that person's group is likely to be viewed negatively and may also be targeted for retaliation because of the individual actor's behavior (Lickel & Onuki, 2015; Lickel, et al. 2006).

Ingroup perspective. From the ingroup's perspective, collective responsibility is reflected in the ingroup's reaction to the ingroup member's behavior that might, or already did, lead to collective blame from the other groups. Ingroup members—fellow group members of the individual actor—respond to actual or possible outgroup blame in different ways. Ingroup members often feel negative emotions when individual group members' actions may bring blame to the group from outsiders (Doosje, Branscombe, Spears, & Manstead, 1998; Lickel, Steele, & Schmader, 2011). For example, many Americans felt negative emotions about the American military occupation of Iraq, particularly when they thought Iraqis viewed Americans' actions as reflecting on America's moral character (Iyer, Schmader, & Lickel, 2007). Ingroup members are thus motivated to keep the ingroup away from situations or behavior that they anticipate would create guilt or shame (Shepherd, Spears, & Manstead, 2013) as well as to see the ingroup behave more positively in order to improve its image and identity (van Leeuwen, 2007; Van Leeuwen, van Dijk, & Kaynak, 2013). Ingroup members might also try to protect the ingroup's reputation by psychologically, symbolically, or factually distancing themselves from the members whose behavior is threatening the ingroup's reputation (Castano, Paladino, Coull, & Yzerbyt, 2002; Eidelman & Biernat, 2003).

Entitativity. Different groups attract collective blame to different degrees. The likelihood of collective blame and ingroup reactions to it depend on the level of the

COLLECTIVE RESPONSIBILITY

perceived interdependence between, and underlying common characteristics of, the group members, which is referred to as the group's *entitativity* (Campbell, 1958; Hamilton & Sherman, 1996). Perceived group entitativity contributes to greater activation of collective responsibility processes both in terms of negative generalization and collective blaming by outgroups (Crawford et al., 2002; Denson et al., 2006; Lickel et al., 2003) and reactions to the individual wrongdoer by the ingroup (e.g., Abrams, Marques, Randsley de Moura, Hutchinson, & Bown, 2004; Lickel et al., 2005).

The work reviewed above testifies to the considerable attention that has been paid to attribution of blame and responsibility at the collective level, as well as people's reaction to wrongdoings of fellow ingroup members. Much less is known about the individual group member whose behavior triggers the collective responsibility processes.

Collective blame: The individual actor's perspective

Given people's propensity to engage in collective responsibility processes both as ingroup and outgroup members, we expected individual actors to be aware that their behavior can reflect on their group and trigger collective responsibility processes. In other words, people should have an intuitive understanding that they bear responsibility for how their group will be seen by outgroup members. Prior research on stereotype threat processes also implies that people are aware of the possibility that individual behavior can affect the ingroup's reputation—at least in terms of its preexisting stereotypes (Cohen & Garcia, 2005; Schmader & Lickel, 2006a). In particular, it has been theorized that people experience group-reputation threat when their behavior could confirm negative stereotypes about their group in the eyes of outgroup members (Shapiro, 2012; Shapiro & Neuberg, 2007). Relatedly, research on helping behavior

COLLECTIVE RESPONSIBILITY

suggests that such group reputation threat may motivate people to behave better. When people felt that the ingroup's image needs improvement (van Leeuwen & Tauber, 2011) or when they were motivated to change an existing negative stereotype about the ingroup (Hopkins, Reicher, Harrison, & Levine, 2007), they helped others more.

We argue, therefore, that people's intuitive understanding of collective responsibility processes (Lickel et al., 2001) also includes the knowledge that their own behavior can trigger these processes. However, we hypothesize that concern about negatively influencing the ingroup's reputation should even occur as a more general phenomenon, without an existing stigma or group stereotype.

People's understanding of their role in triggering collective responsibility processes should manifest in a specific set of *appraisals*, *affective reactions*, and *behavioral responses*. Furthermore, if this understanding is truly an aspect of intergroup psychology, then actors' responses should be sensitive to the context (whether it is an intergroup context or not) and to the nature of the ingroup (whether it is a group that is likely to be viewed by perceivers as an appropriate target for collective blame). Finally, we argue that actors' behavior may be influenced by collective blame processes both before an action (i.e., anticipating collective blame and modifying behavior to prevent it) as well as after an action (i.e., responding when the actor realizes that she/he may have activated collective blame.).

Appraisals. Collective responsibility renders a negative behavior relevant to how the ingroup is perceived by others. People are sensitive to the ingroup's reputation (Branscombe, Ellemers, Spears, & Doosje, 1999; Leach, Ellemers & Barreto, 2007) and tend to see events in light of the likelihood to trigger the ingroup's evaluations by other

COLLECTIVE RESPONSIBILITY

groups (Marques, Abrams, & Serôdio, 2001). Recognizing that their behavior could reflect on the ingroup should increase an actor's concern for the ingroup's reputation. We hypothesized that when collective blame is possible, people are more likely to appraise their own behavior in terms of its potential effects on the ingroup's reputation.

Behavior. People want to view their own group as good or better than other groups and they want other groups to share that view (Tajfel, 1982). Appraising their own behavior in terms of its consequences for the ingroup's reputation should influence behavior. Specifically, understanding that their own behavior may influence the ingroup's reputation should prompt people to behave better. We thus hypothesized that people behave more positively when collective blame is possible.

Emotions. Despite their motivation to protect the group, people may nonetheless engage in actions that can tarnish the group's image either because of insufficient self-control (Baumeister & Heatherton, 1996) or because they realize too late that the behavior could activate collective blame. Re-appraising their past behavior in terms of collective responsibility should elicit negative emotions over and beyond what the person would feel simply for enacting a negative behavior. Further, in order to avoid experiencing such emotions, people should behave more positively in subsequent situations that bear similar consequences. We hypothesized that people would feel more negative emotions about a past behavior if it could activate collective blame, and that the increased negative emotions, in turn, would motivate positive behavior.

Overview of the present research

The above hypotheses were tested across four experiments in which we manipulated the likelihood that a person's individual behavior could activate collective

COLLECTIVE RESPONSIBILITY

blame. We operationalized collective blame possibility by manipulating if the outgroup could make inferences about the individual actor's group. Each experiment was designed to manipulate only the group reputational consequences (collective blame possibility) of one's behavior, while the possible personal consequences of the behavior were kept constant across conditions.

Experiment 1 combined a minimal group paradigm with a behavioral interaction paradigm. Participants played a dictator game and we assessed how prosocially they behaved depending on the possibility that their behavior could influence the ingroup's image, and how they felt about their behavior depending on its likelihood of triggering collective blame. Experiment 2 had participants imagine committing a transgression, manipulating whether or not their group membership became known to outgroup members, and measured ingroup reputation concern and emotional reactions. Using a similar imagined scenario, Experiment 3 manipulated whether ingroup members or outgroup members witnessed the wrongdoing and measured ingroup reputation concern and emotional reactions, while also testing the moderating role of perceived ingroup entitativity. Experiment 4 used a behavioral interaction paradigm manipulating whether or not participants represented the ingroup before an outgroup, and measuring ingroup reputation concern, emotional reactions as well as the emotions' consequences for subsequent behavior.

Statistical power. We performed post-hoc power analyses using the G*Power program (Faul, Erdfelder, Lang, & Buchner, 2007). Based on the obtained effect and sample sizes, the power to detect the effects reported below was .78, .78, .53 and .79 for

COLLECTIVE RESPONSIBILITY

Experiments 1-4, respectively, which, at average, are above the average power of .65 in studies published in upper tier social psychology journals (Fraley & Vazire, 2014).

Experiment 1

We developed a modified dictator game (Camerer, 2003) to test the effect of potential collective blame on actors' behavior. The dictator game is believed to assess prosocial behavior that is not contingent on concern about punishment or explicit desire for reciprocity. In our paradigm, we introduced a reputation system (see Feinberg, Willer, & Schultz, 2014) and manipulated whether individual selfishness could or could not influence the reputation of participants' ingroup in the eyes of outgroup members. Furthermore, to show that actors' sensitivity to collective blame does not depend on intergroup history or group stereotypes, we used a minimal group paradigm (Tajfel, Billig, Bundy, & Flament, 1971). Besides assessing participants' actual behavior, we measured their emotional response, predicting that people who behaved more selfishly would feel worse—but only when their behavior could damage the ingroup's reputation.

Method

Participants. Seventy-eight college undergraduates participated for research credit and for a \$2 monetary reward. Four participants who raised suspicions about the cover story were excluded, leaving seventy-four participants for subsequent analyses ($M_{\text{age}} = 19.89$, $SD = 1.44$, range 18-25; 60 female).

Materials and procedure. According to the cover story, participants played against each other as members of opposing teams (blue vs. red). In reality, there were no other players. Participants first drew a card from a bag to determine which team they were going to join and in which round of the game they would play. Unbeknownst to

COLLECTIVE RESPONSIBILITY

participants, all cards in the bag assigned them to the blue team and to play in game round 1. All participants were told that there were four blue players in the blue group and four red players in the red group; that in each game round, one blue player would play against one red player; and that as “Blue 1,” the participant was to be the first blue player to play (with “Red 1”). The experimenter then led the participant into a room with the sign “Blue 1”. Participants were seated in front of a computer and told to follow the instructions in the computer-assisted survey.

Playing the dictator game. Participants were given detailed instructions as to how to play the game. Participants were told that they were the allocator and “Red 1” the receiver. As allocator, they got 20 aluminum coins, worth 10 cents each. They had to anonymously divide the coins between themselves and “Red 1” as they wished. Participants were told that “Red 1” could not respond to participants and the coins participants owned at the end of the game would be exchanged for real money. To ensure anonymity, a coin acceptor machine was connected to the computer and participants had to insert the coins that they wanted to give to “Red 1.” The coin acceptor would count the coins and send a message to “Red 1,” ostensibly sitting in another room, about how many coins he/she received. The coins that participants wanted to keep had to be placed in an envelope, to be exchanged for real money by a different experimenter later on. Participants were told that after they finish their game with “Red 1,” “Blue 2” would play with “Red 2,” but switching roles so that “Red 2” would be the allocator and “Blue 2” the receiver. Then “Blue 3” would play with “Red 3,” switching roles again, and so on for the fourth pair.

Manipulating collective responsibility in the dictator game. In the collective

COLLECTIVE RESPONSIBILITY

responsibility (but not in the control) condition, participants were also told that at the end of each turn (e.g. after “Blue 1” played “Red 1”), the receivers (e.g. “Red 1”) could send a message to their group members about the allocator’s behavior, where the message indicated the allocator’s group membership (i.e. “Blue 1” was selfish or generous).

Participants’ understanding of the rules was tested with ten questions. After incorrect answers, the survey showed and explained the correct answer. Participants then received the 20 coins, allocated them (alone) and then continued the survey, completing a measure of state emotions.

Emotions. Eight items measured how participants felt after the game: six negative emotions (guilt, remorse, shame, sorry, worry, anxiety), and two positive emotions (satisfaction, happiness) for control purposes and because participants could feel good about earning money. All emotions were rated on continuous scales from 1 (Not at all) to 9 (Very much). Factor analyses suggested two meaningful factors. One of them comprised the six negative emotions ($\alpha = .89$, $M = 2.02$, $SD = 1.31$), the other the two positive emotions ($\alpha = .87$, $M = 5.65$, $SD = 1.72$).

At the end every participant received the maximum possible payout of \$2. After participants were debriefed and left, the experimenter opened the envelopes and recorded how many coins participants had kept for themselves, which formed our dependent variable.

Results

Participants kept significantly fewer coins for themselves (i.e. were more generous) in the collective responsibility ($M = 10.39$) than in the control condition ($M = 12.08$), $F(1, 72) = 7.53$, $p = .008$, $\eta^2 = .09$.

COLLECTIVE RESPONSIBILITY

Next we tested how emotions were influenced by one's behavior depending on the behavior's possible group level consequences. We predicted that the effect of selfish behavior would depend on the presence vs. absence of collective responsibility. We ran a moderated regression with condition as categorical IV and selfish behavior (i.e. number of coins participants had kept for themselves [centered]) as continuous IV, and negative emotions as DV. Selfish behavior had a significant main effect on negative emotions, $F(1, 70) = 10.13, p = .002, \eta^2 = .13$; the more selfishness, the more negative emotions ($\beta = .54$). The main effect of collective responsibility did not reach significance, but showed a trend, $F(1, 70) = 2.68, p = .106, \eta^2 = .04$, with participants reporting more negative emotions when their behavior could reflect on the ingroup ($M = 2.13$) rather than when it could not ($M = 1.95$). Most importantly, these main effects were qualified by the expected two-way interaction between collective responsibility and selfish behavior, $F(1, 70) = 11.64, p = .001, \eta^2 = .14$. Participants who behaved relatively generously (-1SD) reported similar levels of negative emotions when their behavior could reflect on the ingroup ($M = 2.01$) and when it could not ($M = 1.35$), $t(70) = 1.57, p = .12$. Participants who behaved relatively selfishly (+1SD) reported significantly more negative emotions when their behavior could reflect on the ingroup ($M = 3.57$) rather than when it could not ($M = 1.92$), $t(70) = 3.40, p = .001, d = .60$ (see Figure 1). Importantly, and as expected, this interaction was driven by those who behaved selfishly when collective responsibility was possible.

Discussion

As hypothesized, people were more generous to an anonymous outgroup person when their behavior could reflect on their ingroup. In doing so, participants acted against

COLLECTIVE RESPONSIBILITY

their self-interest (earning money) to protect the ingroup's image. Thus, even in minimal groups, concerns about the ingroup's reputation can induce changes in behavior. Also, players who "failed" to heed the potential consequences of their actions for the group and behaved selfishly felt worse in that condition than in the control condition. These two findings highlight that actors' concern for the ingroup's reputation plays a role both before and after action. Experiments 2-4 further unpacked the appraisals, emotions, and future behavior that occur when actors fail to protect the ingroup from collective blame.

Experiment 2

In Experiment 1 participants behaved more positively when their behavior could affect the ingroup's reputation. But, people are not always able to control their behavior. Furthermore, people often realize too late that their actions might have had group level consequences. Collective responsibility processes should nonetheless influence actors' responses to transgressions and influence their behavior afterwards. Experiment 2 tested how people think and feel when they realize that their past behavior could trigger collective blame. We hypothesized that people would then become concerned about the ingroup's reputation and, in turn, feel negative emotions. While Experiment 1 already gave preliminary evidence for this hypothesis, it only did so for people who spontaneously had engaged in selfish behavior that could damage the ingroup's reputation. Experiment 2 instead *manipulated* the framing of participants' past behavior in order to obtain causal evidence.

Method

Participants. 118 Americans were recruited and compensated via Amazon Mechanical Turk ($M_{age} = 35.86$, $SD = 13.53$, range 18-68; 72 female).

COLLECTIVE RESPONSIBILITY

Procedure. Following a widely used method of studying people's reactions to their own wrongdoing, we used imagined scenarios (Baumeister, Stillwell, & Heatherton, 1995). The scenario had participants imagining going on a hiking trip in China, accidentally causing fire damage to a forest, but getting out safely and going home without anybody knowing about their role in the fire. Chinese media then reported on the damage the fire had caused. In the control condition, participants read that Chinese authorities had not been able to gather any information about the person causing the fire, but believed it was a *hiker*. In the experimental condition, Chinese authorities believed it was an *American hiker*. Participants thus damaged the forest in both conditions but in the experimental condition their revealed group membership could additionally damage the ingroup's reputation. All items were measured on continuous scales ranging from 0 (Not at all) to 7 (Very much), unless noted otherwise.

Emotions. Participants completed a self-report measure of state emotionality, rating ten emotions on scales from 0 (Not at all) to 10 (Very intensively). Six items measured negative self-conscious emotions (*guilt, shame, embarrassment, disgrace, remorse, sorry*), two items measured anxiety-related emotions (*worry, anxiety*) and two items measured a positive (*nostalgia*) and a negative (*envy*) emotion for control purposes. Factor analysis led to a three-factor solution. The six negative self-conscious emotions loaded on one factor ($\alpha = .90$, $M = 8.87$, $SD = 1.32$), the two anxiety-related emotions on another ($\alpha = .96$, $M = 8.94$, $SD = 1.54$), and the two unrelated emotions on a third ($\alpha = .65$, $M = 1.01$, $SD = 1.68$). Three univariate outliers on negative self-conscious emotions (2 ½ standard deviations below the mean) were eliminated from further analyses (Tabachnik & Fidell, 2007).¹

COLLECTIVE RESPONSIBILITY

Ingroup reputation concern. Participants appraised the consequences of their behavior answering two questions about the potential effects of the described events on the reputation of the U.S. (“Do you think that the incident described in the scenario may reflect poorly on Americans?” and “Do you think that the incident described in the scenario may somehow damage the image of the United States?”; $\alpha = .87$, $M = 4.89$, $SD = 1.59$).

Appraised cause of emotions. Two items assessed to what extent participants attributed their negative emotions due to the direct consequences of their behavior (i.e. damaging the forest; “I would feel bad because my behavior led to the damage in the national park”, $M = 6.48$, $SD = 0.74$) itself, or to their behavior’s group level consequences (“I would feel bad because my behavior may reflect poorly on other Americans”, $M = 4.50$, $SD = 2.21$).

Attitudes toward China. To control for pre-existing attitudes toward the outgroup, three items assessed participants’ attitudes toward China in general, in cultural and in economic terms ($\alpha = .71$, $M = 3.99$, $SD = 1.18$).

Results

Emotions. A general linear model revealed a significant effect of condition on negative self-conscious emotions, $F(1, 113) = 6.72$, $p = .011$, $\eta^2 = .056$, but not on the anxiety-related or the unrelated emotions, $ps > .10$. People reported significantly more negative self-conscious emotions when they believed that the outgroup was aware of their group membership ($M = 9.16$) than when it was not ($M = 8.62$).

Ingroup reputation concern. Participants’ belief that their actions could hurt the ingroup’s image were significantly stronger in the experimental ($M = 5.27$) than in the

COLLECTIVE RESPONSIBILITY

control condition ($M = 4.54$), $F(1, 116) = 6.58$, $p = .012$, $\eta^2 = .054$.

Appraised cause of emotions. We ran a repeated-measures analysis of variance (ANOVA) with condition as independent variable, the two appraised causes (damaging the forest and damaging U.S. reputation) as dependent variables, and the type of cause as repeated factor. The condition by type interaction was significant, $F(1, 116) = 5.71$, $p = .018$. Participants' attributions of their negative self-conscious emotions to damaging the forest did not significantly differ between conditions, $F(1, 116) = .97$, $p = .326$, whereas their attributions of their negative self-conscious emotions to harming U.S. reputation were significantly higher in the experimental ($M = 5.09$) than in the control condition ($M = 3.96$), $F(1, 116) = 8.08$, $p = .005$, $\eta^2 = .065$. This finding indicated that the condition effect on emotions was carried by the concern over damage to the ingroup's reputation, not by the concern over the damage to the outgroup and its resources (i.e. the Chinese forest).

Mediational analysis. In emotion theory, the appraisal of the event is conceptualized as the trigger of the emotion (Frijda, Kuipers, & ter Schure, 1989). We thus tested whether concern over the ingroup's reputation mediated the effect of condition on negative self-conscious emotions, using 95% bias-corrected confidence intervals and 5000 bootstrap resamples (Hayes, 2012, model 4). In this model, condition significantly predicted ingroup reputation concern, $t = 2.63$, $p < 0.01$, which in turn predicted negative self-conscious emotions, $t = 2.45$, $p = .016$. The confidence interval of the indirect effect did not include zero ($CI_{95} = [.021, .318]$), indicating a significant indirect effect (see Figure 2).

Discussion

COLLECTIVE RESPONSIBILITY

In line with our hypothesis, imagining oneself as the perpetrator of a negative action led to more negative emotions when the outgroup learned about the perpetrator's group membership. Participants felt negative emotions for a harmful behavior in both conditions, but this response was significantly stronger when the behavior could also damage the ingroup's reputation. Furthermore, as predicted, the perpetrator's concern over the ingroup's reputation mediated the effect of collective blame possibility on negative emotions. These results lend direct support to the hypothesis that collective responsibility processes affect the individual actor even *after* an action has already happened.

Experiment 3

Past research on collective responsibility has found that people are more likely to blame groups high (rather than low) in entitativity (Denson, et al., 2006; Lickel, et al. 2003). The reason is that one has to assume a certain level of similarity and connection among the members of a group to judge the whole group based on one of its members' behavior. If the wrongdoer's ingroup is less entitative, its members are less concerned about drawing collective blame (Lickel, Schmader, & Barquissau, 2004). We expected that this logic would extend to the individual actor. Only when the ingroup is perceived as highly entitative should the wrongdoer expect that outsiders are likely to generalize the behavior to the wrongdoer's ingroup. This possibility of collective blame should then increase the reputation concern and negative emotions that the wrongdoer experiences over the wrongdoing. Testing for this expected interaction of collective blame possibility and entitativity, Experiment 3 also addressed a possible confound of Experiment 2. In Experiment 2, the wrongdoing happened abroad in both conditions, possibly priming an

COLLECTIVE RESPONSIBILITY

intergroup context even in the control condition. While this would, if anything, have worked against the effect we predicted, Experiment 3 avoided this possible confound by manipulating whether the same transgression happened at home or abroad, thus establishing better experimental control over whether the context was intergroup or intragroup. Experiment 3 further measured the wrongdoer's concern about both the *ingroup's* and his or her *personal* image in order to distinguish between these two types of image (concern), and used a less blatant measure of the perceived damage to the ingroup's image than Experiment 2.

Method

Participants. Of 208 Americans recruited and paid via Amazon Mechanical Turk, eighteen could not correctly recall information from the manipulation material during the manipulation check and were thus excluded from further analysis; 190 participants remained ($M_{age} = 34.96$, $SD = 13.43$, range 18-73; 108 female).

Materials and Procedure. In both conditions participants imagined themselves as the main character in a scenario. The scenario described two friends getting into a heated argument on the street, observed by local bystanders. In one condition it happened during a vacation in a Southern Californian city (intragroup condition), whereas in the other condition it happened in an unspecified foreign country (intergroup condition). Therefore, the bystanders witnessing the quarrel were strangers in both conditions, but only in the intergroup condition were the witnesses members of a national outgroup who could form a negative opinion about participants' national ingroup.

Emotions. After reading the scenario, participants completed an emotion response scale, rating nine emotions on continuous scales from 0 (Not at all) to 10 (Extremely

COLLECTIVE RESPONSIBILITY

much). As in Experiment 3, we measured six negative self-conscious emotions (*guilt, shame, embarrassment, disgrace, remorse, sorry*) and two anxiety-related emotions (*worry, anxiety*). The ninth emotion (*envy*) was used as a control item². Factor analysis suggested one meaningful factor that explained 95% of the variance of the items. Eight items loaded on this factor, which we labeled *negative emotions* ($\alpha = .90$, $M = 6.56$, $SD = 1.82$).³ Envy ($M = 1.26$, $SD = 1.52$) loaded below .40 on this factor and was thus dropped.

Ingroup reputation concern. A subtle measure of ingroup reputation concern followed, consisting of five items in total. The first item was open-ended, asking participants why they would feel the reported emotions. Three more items asked participants to complete unfinished sentences: “People who saw the incident may form negative opinions about ...”, “My behavior may reflect poorly on ...”, “I would feel responsible for possibly causing troubles to...”. Finally, the fifth item was again open-ended (“What kind of negative effects might your behavior in the story have had?”). All items could be answered referring to the self or the ingroup. Two independent coders judged whether or not participants referred to the U.S. (ingroup) in their answers. If participants referred to the U.S., it was coded 1, if not it was coded zero. The sum of the resulting five codings constituted an index of participants’ concern over the ingroup ($M = 0.91$, $SD = 1.40$). Participants then also answered the same ingroup reputation concern questions as in Experiment 2 ($\alpha = .90$, $M = 3.56$, $SD = 2.02$). This explicit measure of ingroup reputation concern strongly correlated with the subtle measure, $r = .52$, $p < .001$ and factor analysis showed that the items of the two measures loaded onto the same

COLLECTIVE RESPONSIBILITY

factor. The explicit and subtle scores were thus each standardized and averaged into a composite score of ingroup reputation concern ($M = 0.00$ and $SD = 0.87$).

Appraised cause of emotions. We again assessed emotion attribution with two items, but here we focused on parallel assessment of the transgression's consequences for the ingroup's reputation ("I would feel bad because my behavior may reflect poorly on other Americans"; $M = 3.72$, $SD = 2.42$) and for the individual reputation ("I would feel bad because my behavior may reflect poorly on me"; $M = 5.96$, $SD = 1.28$). Participants' answers were recorded on continuous scales from 0 (Not at all) to 7 (Very much).

Perceived ingroup entitativity. Perceived entitativity of the U.S. was assessed with five items adapted from past research (e.g. "How cohesive of a group is the United States?") (Lickel et al., 2003). The answers were rated on continuous scales from 0 (Not at all) to 7 (Very much), $\alpha = .81$, $M = 3.96$, $SD = 1.19$.

Manipulation check. Participants had to recall whether the location of the argument was in a foreign country or in Southern California.

Results

We ran moderated regression analyses with condition as categorical IV, entitativity (centered; not affected by condition, $p = .206$) as a continuous moderator, and with negative emotions, ingroup reputation concern and attributed cause of emotions as DVs.

Emotions. The main effect of condition on negative emotions was marginally significant, $F(1, 186) = 2.83$, $p = .094$, $\eta^2 = .015$. Negative emotions were lower in the intragroup ($M = 6.37$) than in the intergroup condition ($M = 6.72$), consistent with the finding of Experiment 2. Higher entitativity predicted more negative emotions, $F(1, 186) = 9.58$, $p = .002$, $\eta^2 = .049$, $\beta = .41$. Most importantly, the predicted interaction between

COLLECTIVE RESPONSIBILITY

entitativity and condition was significant, $F(1, 186) = 5.85, p = .017, \eta^2 = .031$.

Entitativity significantly influenced negative emotions in the intergroup condition, $t(186) = 4.67, p < .001$, but not in the intragroup condition, $t(186) = .042, p = .676$. From another angle, participants who perceived the U.S. as high in entitativity (+1 SD) felt significantly more negative emotions when collective responsibility was present ($M = 7.52$) than when it was absent ($M = 6.45$), $t(186) = 2.96, p = .004, d = .43$. Participants who viewed the U.S. as low in entitativity (-1SD), on the other hand, felt similar levels of negative emotions when collective responsibility was present ($M = 6.06$) or absent ($M = 6.28$), $t(186) = -.58, p = .562$ (see Figure 3).

Ingroup reputation concern. Analyses showed significant main effects of condition, and entitativity. Group reputation concern was significantly higher in the intergroup ($M = 0.56$) than in the intragroup condition ($M = -0.61$), $F(1, 186) = 174.17, p < .001, \eta^2 = .48$, and higher entitativity predicted greater perceived damage to the ingroup's image ($\beta = .16$), $F(1, 186) = 11.22, p = .001$. The interaction effect did not reach significance but showed a trend, $F(1, 186) = 2.70, p = .102, \eta^2 = .014$, and the simple effects were in line with predictions. Entitativity significantly influenced concern in the intergroup condition, $t(186) = 4.23, p < .001$, but not in the intragroup condition, $t(186) = 1.06, p = .292$. From another angle, participants who perceived the U.S. as high in entitativity (+1 SD) were significantly more concerned when collective responsibility was present ($M = .81$) rather than when it was absent ($M = -.54$), $t(186) = 10.47, p < .001$. Participants who viewed the U.S. as low in entitativity (-1SD) were also more concerned when collective responsibility was present ($M = .34$) rather than when it was absent ($M = -.70$), but to a lesser extent, $t(186) = 7.78, p < .001$.

COLLECTIVE RESPONSIBILITY

Attributed cause of emotions. People's attributions of their emotions to concern over damage of their *personal* image did not differ based on condition, entitativity, or their interaction, $p_s > .149$. People's attributions of their emotions to concern over damage of their *group's* image, on the other hand, were significantly higher in the intergroup ($M = 4.63$) than in the intragroup condition ($M = 2.70$), $F(1, 186) = 45.45, p < .001, \eta^2 = .195$, and they were positively predicted by entitativity ($\beta = .66$), $F(1, 186) = 17.19, p < .001, \eta^2 = .085$. Importantly, these effects were qualified by a marginally significant two-way interaction between condition and entitativity, $F(1, 186) = 2.77, p = .098, \eta^2 = .015$. In the intragroup condition, entitativity did not predict attributions of emotions to concern over damage of the ingroup's image, $t = 1.54, p = .125$, whereas in the intergroup condition it did, $t(186) = 4.92, p < .001, d = .72$. Participants who perceived the U.S. as high in entitativity (+1 SD) attributed their emotions significantly more to concern over damage of the ingroup's image when collective responsibility was present ($M = 5.64$) rather than when it was absent ($M = 3.06$), $t(186) = 5.94, p < .001$. Participants who viewed the U.S. as low in entitativity (-1SD) also attributed their emotions significantly more to concern over damage of the ingroup's image when collective responsibility was present ($M = 3.78$) rather than when it was absent ($M = 2.26$), but to a lesser extent, $t(186) = 3.37, p < .001$.

Mediational analyses. We conducted a moderated mediation analysis, with condition as IV, entitativity as a moderator, ingroup reputation concern as mediator, and negative emotions as the DV, with 95% bias-corrected confidence intervals and 5000 bootstrap resamples (Hayes, 2012, model 59). The confidence intervals of the indirect effect included zero ($CI_{95} = [-.048, .526]$) at low levels of entitativity, but not at high

COLLECTIVE RESPONSIBILITY

levels of entitativity ($CI_{95} = [.022, .863]$), indicating that ingroup reputation concern carried the condition effect among those who perceived the U.S. as highly entitative.⁴

Discussion

The results conceptually replicated the findings of Experiment 2 with a different manipulation and with a new outcome variable. Most importantly, they confirmed our hypothesis that entitativity moderates collective responsibility processes within the individual actor as well. Even when collective blame was possible, only wrongdoers who perceived the ingroup as highly entitative felt more negative emotions for a past wrongdoing. Importantly, participants' attributions of their emotions to concern over damage of their personal image were similar in both conditions, but their attributions to concern over damage of the *ingroup's* image were significantly higher in the intergroup condition. Also, similar to Experiment 2, the condition effect on negative emotions was mediated by ingroup reputation concern.

In Experiment 1, we showed that people regulate their (present) behavior when it can impact the ingroup's reputation. Experiment 2 and 3 showed that people feel bad when their past behavior (could have) triggered collective blame, and that this effect was mediated by ingroup reputation concern and moderated by entitativity. Experiment 4 tested whether the negative emotions people experience after triggering collective blame motivate them to behave better in the *future*.

Experiment 4

Experiments 2 and 3 demonstrated that people feel more negatively about past wrongdoings if they could activate collective blame of the ingroup. We hypothesized that this after-the-fact response is not merely emotional, but should also influence subsequent

COLLECTIVE RESPONSIBILITY

behavior to better protect the ingroup's reputation. Experiment 4 tested this hypothesis with a modified, two-round dictator game paradigm. In doing so, Experiment 4 served as an important replication of Experiments 2 and 3, while also extending the former studies' use of vignettes to actual behavior. Specifically, we expected that people would experience more negative emotions when their alleged selfishness (manipulated with bogus feedback) in the first game round could hurt the ingroup's reputation, and that this increase in negative emotions would lead to less selfish behavior in the second game round (see Figure 4).

Method

Participants. Seventy-seven college undergraduates ($M_{\text{age}} = 19.73$, $SD = 1.26$, range 18-23 years; 66 female) participated for research credit and a \$2 monetary reward.

Materials and procedure.

Participants first completed a measure of perceived entitativity of their university's student body.

Entitativity. We adapted five items from a perceived group entitativity scale (Ip, Chiu, & Wan's, 2006) tapping the ingroup's perceived cohesiveness (e.g. "To what extent you think that others perceive *the university's* students as a cohesive real group?") and homogeneity (e.g. "Overall, how similar do you think are *the university's* students to each other?"). All items were measured on continuous scales from 1 (Strongly disagree) to 9 (Strongly agree) and loaded on one factor ($\alpha = .69$, $M = 5.45$, $SD = 0.93$).

Dictator game. Participants played the game as the allocator, with another player ostensibly sitting in another room; in reality, there was no other player. After participants read the instructions, the experimenter tested for their understanding of the game and

COLLECTIVE RESPONSIBILITY

walked them through the rules again. At this point, participants got twenty aluminum coins, each worth five cents, to anonymously divide between themselves and the other player. Participants were informed that there was no danger of retaliation from the other player, and that at the end of the study they would get real money for the coins they had kept for themselves. The coins that participants kept for themselves had to be inserted into a coin acceptor machine connected to the computer. The coins that participants wanted to give to the other player had to be placed into an envelope. Participants were told that the computer automatically counted the coins inserted into the coin acceptor and therefore their decision would remain unknown to the experimenter. The experimenter then left the room and participants distributed the coins. Then the experimenter returned, took the envelope and ostensibly brought it to the other player. In reality, outside the testing room the experimenter labeled and filed the envelope for data entry. Participants meanwhile continued the survey.

Induction of negative emotions. The survey instructions informed participants in both conditions that the game that they had just played measures moral character based on how many coins they had given to the other player. They were told that the computer had calculated that they gave four coins less to the other player than others gave on average. This average remained unspecified, so that almost everybody could believe this feedback. Only participants who had given more than 16 of their 20 coins should find the feedback illogical and confusing, and thus would need to be excluded from later analyses, but a meta-analysis found that only 6.5% of people give away more than 80% of their endowment in dictator games (Engel, 2010). Importantly, as all participants received the same feedback, the moral transgression and the possible emotional consequences of the

COLLECTIVE RESPONSIBILITY

transgression per se were kept constant across conditions.

Manipulation of collective responsibility. To manipulate collective responsibility, in the experimental condition the instructions additionally mentioned that the study was part of an inter-college comparison assessing students' moral character in the region. They were told that the computer had randomly selected their data to compare the two in-town colleges in moral character. To play down any personal accountability, participants were assured that their individual/personal identity was kept completely anonymous and their personal names would not appear in any report of the data. Their college's name, however, would appear next to their data. Thus, while participants in both conditions were led to believe that they had behaved selfishly, only participants in the experimental condition were further led to believe that this individual selfish behavior could also reflect on their ingroup.

Emotions. After the first game, participants reported how they felt. Eight items measured negative emotions (*guilt, remorse, disgrace, shame, sorry, worry, anxiety, anger*⁵), and four items measured positive emotions (*pleasure, satisfaction, contentment, joy*). All emotions were rated on continuous scales from 1 (Not at all) to 9 (Very much). Factor analysis suggested two meaningful factors, one comprising the eight negative items (*round 1 negative emotions*; $\alpha = .94$, $M = 4.05$, $SD = 1.70$), the other the four positive emotions (*round 1 satisfaction*; $\alpha = .79$, $M = 4.36$, $SD = 1.41$).

Ingroup reputation concern. On continuous scales from 1 (Not at all) to 9 (Very much), participants expressed their agreement with four statements assessing their appraisal of the consequences of their actions for the ingroup's reputation (e.g. "I am afraid that my decision in the game damaged <participants' university's> reputation", "I

COLLECTIVE RESPONSIBILITY

worry that the way I played the game could paint a poor image about <participants' university> students' moral character"). All items formed one factor, labeled *round 1 ingroup reputation concern* ($\alpha = .97$, $M = 3.31$, $SD = 2.03$).

Game round 2. Participants were then informed that they would play the game again, and again as the allocator. They were walked through the rules again and got a new set of twenty coins. After distributing the coins between themselves and the other player (the same way as in game round 1), participants continued the survey. They were told that they would receive feedback about their performance in the second game later; in reality, this feedback never came. Participants then completed the same emotion and ingroup reputation concern measures they had completed after the first round; factor analyses suggested the same factors as for the round 1 measures. Composite scores were computed by averaging the eight negative emotion items into *round 2 negative emotions* ($\alpha = .92$, $M = 2.90$, $SD = 1.41$), the four positive emotions into *round 2 satisfaction* ($\alpha = .73$, $M = 5.32$, $SD = 1.11$), and the ingroup reputation concern items into *round 2 ingroup reputation concern* ($\alpha = .97$, $M = 3.57$, $SD = 2.25$).

At the end of the study, participants completed basic demographics and were asked to describe the study's goal. All participants received the maximum possible payout of \$2 and were debriefed. The experimenter opened the envelopes and recorded the number of coins participants had given to the other player in the first envelope (*round 1 coins*) and in the second envelope (*round 2 coins*) as (behavioral) dependent variables.

Results

Consistent with the aforementioned meta-analysis, in our study only seven out of 77 participants (9.1%) gave more than 16 coins. Since the bogus feedback was

COLLECTIVE RESPONSIBILITY

nonsensical for them, they were excluded from analyses. Further, four participants did not follow the instructions (one failed to notify the experimenter when prompted, and three terminated the study prematurely), and three participants suspected that there was no other player. It is not uncommon that some participants are suspicious of the authenticity of bogus feedback, or of the existence of bogus partners (e.g. Galinsky, Magee, Gruenfeld, Whitson, & Liljenquist, 2008; Muller & Butera, 2007). Sixty-three participants remained for subsequent analysis.

Although entitativity had been measured before the manipulation, and participants had been randomly assigned to conditions, entitativity was higher in the experimental ($M = 5.83$) than in the control condition ($M = 5.08$), $F(1, 61) = 11.80, p = .001$. To make sure that this failure of random assignment did not account for other results, we performed all analyses reported below also with entitativity as a covariate; all remained essentially unchanged.

Effects of collective responsibility on ingroup reputation concern and negative emotions. To answer our primary question of how collective responsibility affects ingroup reputation concern and negative emotions, we introduced ingroup reputation concern and negative emotions as DVs and condition as IV into general linear models. As predicted, *round 1 ingroup reputation concern* was significantly higher in the experimental ($M = 4.46$) than in the control condition ($M = 2.19$), $F(1, 61) = 28.26, p < .001, \eta^2 = .32$. *Round 2 ingroup reputation concern* was also higher in the experimental ($M = 4.93$) than in the control condition ($M = 2.26$), $F(1, 61) = 34.19, p < .001, \eta^2 = .36$. Similarly, *round 1 negative emotions* were significantly higher in the experimental ($M = 4.56$) than in the control condition ($M = 3.56$), $F(1, 61) = 5.77, p = .019, \eta^2 = .09$. *Round*

COLLECTIVE RESPONSIBILITY

2 *negative emotions* were also significantly higher in the experimental ($M = 3.44$) than in the control condition ($M = 2.42$), $F(1, 61) = 8.36$, $p = .005$, $\eta^2 = .12$. Neither round 1 nor round 2 *satisfaction* differed between conditions, $F_s < .33$, $p_s > .56$.

The mediating role of ingroup reputation concern. We tested whether perceived ingroup reputation concern mediated the effect of condition on negative emotions, using 95% bias-corrected confidence intervals and 5000 bootstrap resamples (Hayes, 2012, model 4). We introduced condition as IV, round 1 ingroup reputation concern as mediator, and round 1 negative emotions as DV. Condition significantly predicted round 1 ingroup reputation concern, $t = 5.31$, $p < .001$, which in turn predicted round 1 negative emotions, $t = 4.84$, $p < .001$. The indirect effect was significant ($CI_{95} = [.368, 1.042]$). Unlike in Experiment 1, where the number of coins predicted negative emotions, in Experiment 4 we did not expect that round 1 coins would predict negative emotions because we gave every participant the same (bogus) feedback about how selfishly they had behaved. Accordingly, the number of coins that participants gave to the other player did not, and should not, predict negative emotions or ingroup reputation concern after the first game.

Effects of ingroup reputation concern and negative emotions on future behavior. Next we tested how collective responsibility, ingroup reputation concern and negative emotions after game 1 influenced participants' behavior in game 2. First we ran a repeated-measures ANOVA with the number of coins people gave away in the first game (round 1 coins) and in the second game (round 2 coins) as DVs, condition as IV, and game round as the repeated-measures factor. The condition by game round interaction was significant, $F(2, 61) = 5.91$, $p = .018$, indicating that the effect of condition on

COLLECTIVE RESPONSIBILITY

behavior in round 1 was different from the effect of condition on behavior in round 2.

The subsequent univariate analyses showed that, as predicted, only *round 2 coins* was significantly affected by condition, $F(1, 61) = 6.43, p = .014, \eta^2 = .10$, but not *round 1 coins*, $F(1, 61) = 2.71, p = .105, \eta^2 = .04$.

The mediating role of negative emotions. To fully uncover how collective responsibility shapes subsequent behavior and to see what role emotions play in it, we ran a mediational analysis with 95% bias-corrected confidence intervals and 5000 bootstrap resamples (Hayes, 2012, model 4) with condition as IV, round 1 negative emotions as mediator, round 2 coins as DV, and round 1 coins as covariate. Condition significantly predicted round 1 negative emotions, $t = 2.53, p = .014$, which, in turn, marginally significantly predicted round 2 coins, $t = 1.85, p = .069$. The indirect effect was significant ($CI_{95} = [.064, 1.00]$), and the direct effect became non-significant, $t = 1.24, p = .22$.

The effect of future behavior on subsequent negative emotions. Finally, we tested whether the more positive behavior in the second game would alleviate negative emotions after the second game. We ran a sequential mediation (Hayes, 2012, model 6) with condition as IV, round 1 negative emotions and round 2 coins as mediators, round 2 negative emotions as DV, and round 1 coins as covariate. This analysis reflects the hypothesized process, in which condition (collective responsibility) should influence the negative emotions felt after the first game, which in turn should influence how many coins people gave away in the second game—a behavior that should eventually alleviate negative emotions (i.e. reduce negative emotions felt after the second game). Supporting our hypothesis, condition significantly influenced round 1 negative emotions, $t = 2.52, p$

COLLECTIVE RESPONSIBILITY

= .014, which, in turn, marginally significantly influenced round 2 coins, $t = 1.85$, $p = .069$. Round 2 coins, in turn, marginally significantly influenced round 2 negative emotions, $t = -1.79$, $p = .079$. The sequential indirect effect (condition > round 1 negative emotions > round 2 coins > round 2 negative emotions) was significant ($CI_{95} = [-.127, -.001]$; see Figure 5), and the direct effect became non-significant, $t = 1.40$, $p = .166$. Of note, as predicted, in the collective blame condition negative emotions about one's selfish behavior predicted more generous behavior in the next game, and this more generous behavior then alleviated negative emotions. In the alternative models that did not include round 1 negative emotions or that reversed the order of round 1 negative emotions and round 2 negative emotions, indirect effects were not significant.

Discussion

Experiment 4 replicated the findings of Experiment 2 and 3 in the laboratory, using an economic game paradigm. Participants felt more negative emotions when they learned (after the fact) that their past, *actual* behavior (in the first game) was selfish and thus could have caused collective blame for the ingroup. This increase in negative emotions led participants to behave more generously in the second game. In turn, this more generous behavior alleviated negative emotions. Importantly, participants remained anonymous and thus there was no threat to their individual reputation. The consequences of negative emotions (in response to past behavior) for future behavior shown in this study also rendered additional evidence that the reported emotions were indeed experienced, allaying potential concerns over demand characteristics.

General Discussion

By synthesizing the ingroup and outgroup perspectives of collective

COLLECTIVE RESPONSIBILITY

responsibility, we concluded that there is a third perspective that had yet to be understood. This third perspective is that of the individual actor's, whose behavior triggers collective blame. Four experiments yielded converging evidence that collective responsibility processes influence the individual actor as well. When collective blame was possible, people behaved more prosocially (Experiment 1) – even against their own personal interest. Similarly, for past behavior, when collective blame was possible people appraised their behavior in terms of its (potential) consequences for the ingroup's reputation. People experienced more ingroup reputation concern and negative emotions for the same past behavior when they thought that it could reflect negatively on the ingroup in the eyes of outgroups (Experiment 2-4). This increase in negative emotions further led to subsequent behavior aimed to improve the ingroup's image, even at the actor's personal expense (Experiment 4). These effects of collective responsibility on appraisals and emotions were further moderated by perceived ingroup entitativity (Experiment 3). The fact that in all three perspectives of collective responsibility (i.e., outgroup, ingroup, actor) entitativity serves as a moderator, corroborates our contention that the seemingly distinct psychological mechanisms in each perspective indeed reflect the *same* broader psychological phenomenon—that of collective responsibility. The converging evidence for our proposed third perspective of collective responsibility and its underlying psychological processes is robust given that the four experiments reported here used different paradigms (imagined responses to a vignette vs. behavior in experimental games), experimental settings and data collection strategies (online vs. laboratory), as well as types of groups (national vs. college identity vs. minimal group).

The self-regulatory function of collective responsibility and the intergroup context

COLLECTIVE RESPONSIBILITY

The individual actor's perspective also reveals how collective responsibility processes can regulate individual behavior. One of the responsibilities of a good group member is to not bring blame and disrepute to the group. In this sense, collective responsibility can be viewed as an additional source of *self-regulation*, where anticipated or actual group level consequences guide individual behavior. Self-regulation is generally seen as the basis of proper conduct and peaceful social life (Tangney, Baumeister, & Boone, 2004). Recognizing the possibility to damage the ingroup's reputation and the motivation to avoid such damage may provide additional regulatory resources to think and act properly.

Remarkably, collective responsibility exercises its self-regulatory power specifically in *intergroup* situations—an area of social life that is widely seen as being prone to destructive behavior. Intergroup contexts have frequently been conceptualized as a source of animosity, conflict, and violence that, compared to intragroup contexts, has detrimental effects on behavior (e.g., Wildschut, Pinter, Vevea, Insko, & Schopler, 2003). However, in line with an emerging interpretation of the intergroup context as a possible source of positive effects (e.g. Spears, 2010), we have shown that the intergroup context can promote positive behavior through collective responsibility processes. For the individual actor, intergroup contexts make salient the responsibility of being a good representative of the ingroup. This role as a group representative requires individual group members to manage the impression of their ingroups, extending the scope of impression management from personal to social self, and from personal to group reputation.

The feeling of hurting the ingroup's image

COLLECTIVE RESPONSIBILITY

We found that people experience negative emotions for tarnishing the ingroup's image and that these emotions in turn motivate positive behavior that can improve the ingroup's image. We measured emotions related to guilt and shame, and in each study they formed one single factor. This is consistent with research showing that although guilt and shame are distinct (e.g., Tangney, Miller, Flicker, & Barlow, 1996), they do correlate (e.g., Schmader & Lickel, 2006b) and have many underlying appraisals in common (Smith & Ellsworth, 1985). Of importance in our focus on prosocial behavior and intergroup contexts, both guilt and shame can motivate compensatory behavior after intergroup wrongdoing (Brown, González, Zagefka, Manzi, & Cehajic, 2008; Gausel, Leach, Vignoles, & Brown, 2012). In this way, ingroup members' reactions to collective blame in terms of collective guilt and shame motivate compensatory behavior *indirectly*, through the wrongdoer's ingroup members. It is possible that compensatory motives also play a role in the individual actor's emotional response to triggering collective blame. Complementing this finding, our results suggest that collective responsibility also motivates compensation *directly*, through the wrongdoer him- or herself, for example when in our Experiment 4 negative emotions after game 1 led to more positive behavior in game 2. While we have only tested the effects of negative emotions in general, future research should tease apart possibly differential effects of guilt over and shame for bringing collective blame to one's group (Deonna & Teroni, 2008; Gausel & Leach, 2011).

Boundary conditions and future directions

In the present set of studies, participants were aware that their individual behavior could hurt the ingroup's reputation. In real life, however, the relationship between

COLLECTIVE RESPONSIBILITY

individual behavior and its group level consequences can be ambiguous, and thus the right course of action uncertain. One reason for this ambiguity is that people's judgments of what is good or bad for the ingroup are often inaccurate, and their subjective estimates of how much their personal opinion is representative of their group are often inflated (Haslam, Oakes, McGarty, Turner, & Onorato, 1995). Being convinced that the ingroup should display toughness, for example, individual actors may engage in violence out of the false belief of serving the ingroup's ideal reputation. Ingroup heterogeneity and the existence of subgroups may also contribute to ambiguity. Based on their political beliefs, for example, American liberals and conservatives may perceive different behaviors as suitable for promoting a positive image of the U.S. in the world.

Another factor that might limit prosocial behavior in intergroup contexts is that of pre-existing (perceived) intergroup hostility. If an individual group member believes the ingroup or some of its members to be hostile toward the outgroup, the individual may not engage in prosocial behavior in intergroup contexts, or even engage in antisocial behavior, as the individual may not be concerned about the ingroup's image in the eyes of a hostile outgroup. In fact, rather than behaving more prosocially, in situations characterized by intergroup hostility people may purposefully provoke negative group level consequences. Also, in many situations the actor does not act alone and is perceived as an individual representative of the ingroup but, instead, acts together with other members following the directives and norms of the ingroup.

Finally, the individual actor's perspective connects with research on stereotype threat. The current research shows that people's awareness that their individual behavior can influence the ingroup's reputation goes beyond their awareness that their individual

COLLECTIVE RESPONSIBILITY

behavior can confirm negative group stereotypes. While we expect that the actor's concern over collective blame would likely be strengthened if their actions could confirm a pre-existing negative stereotype of the group, our findings (particularly using minimal groups in Experiment 2) suggest that such stereotype threat is not *necessary* for collective responsibility processes at the individual level to occur. These findings also indicate that group reputation concern (Shapiro & Neuberg, 2007) may be a more general phenomenon than previously believed, affecting behavior in many intergroup contexts rather than only those characterized by pre-existing group stereotypes.

Concluding remarks

Psychology has long been investigating collective responsibility, showing that individual behavior has group level consequences. Similarly, psychology has long been investigating the intergroup context, showing that it often has negative consequences for behavior towards outgroup members. Introducing a new perspective of collective responsibility, that of the individual actor who activates collective responsibility processes, we have shown that individual actors are aware of the group level consequences of collective responsibility, and that in this case, the intergroup context can have positive consequences for behavior towards outgroup members. In intergroup contexts where individual behavior can trigger collective blame, the individual tries to not bring blame and disrepute to the ingroup by behaving more prosocially, even toward outgroup members. When the individual realizes he or she failed to do so (inadvertently or not), he or she becomes concerned, feels bad, and tries to remedy it. By learning from this new perspective how collective responsibility processes guide individual behavior, we may be able to leverage collective responsibility to improve people's behavior even in

COLLECTIVE RESPONSIBILITY

intergroup contexts and towards outgroup members.

COLLECTIVE RESPONSIBILITY

Notes

¹ In all of the studies we measured emotions first, before measuring ingroup reputation concern. Asking first about ingroup reputation concern could have distorted responses and possibly led to demand characteristics on the emotion measure.

² We used the same list of emotions as in Experiment 3 except dropping one of the control emotions (nostalgia).

³ Using the same 6-item negative emotions factor as in Experiment 3 produced virtually the same results in every analysis.

⁴ We also replicated the simple mediation as in Experiment 3.

⁵ Compared to Experiment 4, we replaced embarrassment with anger.

COLLECTIVE RESPONSIBILITY

References

- Abrams, D., Marques, J. M., Randsley de Moura, G., Hutchinson, P. & Bown, N. J. (2004). The maintenance of entitativity: A subjective group dynamics approach. In V. Y. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: contributions to the study of homogeneity, entitativity, and essentialism* (pp. 361-380). Philadelphia, PA: Psychology Press.
- Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry*, 7, 1-15.
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1995). Personal narratives about guilt: Role in action control and interpersonal relationships. *Basic and Applied Social Psychology*, 17, 173-198.
- Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 35-58). Oxford, UK: Blackwell.
- Brown, R., González, R., Zagefka, H., Manzi, J., & Cehajic, S. (2008). Nuestra culpa: Collective guilt and shame as predictors of reparation for historical wrongdoing. *Journal of Personality and Social Psychology*, 94, 75-90.
- Camerer, C. (2003). *Behavioral Game Theory. Experiments in Strategic Interaction*. New York, NY: Sage.
- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, 3, 14-25.

COLLECTIVE RESPONSIBILITY

- Castano, E., Paladino, M., Coull, A., & Yzerbyt, V. Y. (2002). Protecting the ingroup stereotype: Ingroup identification and the management of deviant ingroup members. *British Journal of Social Psychology, 41*, 365–385.
- Cohen, G. L., & Garcia, J. (2005). "I am us": Negative stereotypes as collective threats. *Journal of Personality and Social Psychology, 89*, 566-582.
- Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entitativity, stereotype formation, and the interchangeability of group members. *Journal of Personality and Social Psychology, 83*, 1076-1094.
- Denson, T. F., Lickel, B., Curtis, M., Stenstrom, D. M., & Ames, D. R. (2006). The roles of entitativity and essentiality in judgments of collective responsibility. *Group Processes & Intergroup Relations, 9*, 43-61.
- Deonna, J. A. & Teroni, F. (2008). Differentiating shame from guilt. *Consciousness and Cognition, 17*, 1063-1400.
- Doojse, 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.
- Eidelman, S., & Biernat, M. (2003). Derogating black sheep: Individual or group protection? *Journal of Experimental Social Psychology, 39*, 602–609.
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics, 14*, 583-610.
- 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.

COLLECTIVE RESPONSIBILITY

- Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and Ostracism Promote Cooperation in Groups. *Psychological Science*, 25, 656-664.
- Fiske, S. T. & Taylor, S. E. (2013). *Social Cognition: From Brains to Culture*. London: Sage.
- Fraley, R. C., & Vazire, S. (2014). The N-pact factor: Evaluating the quality of empirical journals with respect to sample size and statistical power. *PLoS ONE*, 9(10): e109019.
- Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal and emotional action readiness. *Journal of Personality and Social Psychology*, 57, 212–228.
- Galinsky, A. D., Magee, J. C., Gruenfeld, D. H, Whitson, J., & Liljenquist, K. A. (2008). Social power reduces the strength of the situation: Implications for creativity, conformity, and dissonance. *Journal of Personality and Social Psychology*, 95, 1450- 1466.
- Gausel, N. & Leach, C. W. (2011). Concern for self-image and social-image in the management of moral failure: Rethinking shame. *European Journal of Social Psychology*, 41, 468-478.
- Gausel, N., Leach, C. W., Vignoles, V. L., & Brown, R. J. (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, 941-960.
- Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review*, 103, 336-355.

COLLECTIVE RESPONSIBILITY

- Haslam, S. A., Oakes, P. J., McGarty, C., Turner, J. C., & Onorato, S. (1995). Contextual changes in the prototypicality of extreme and moderate outgroup members. *European Journal of Social Psychology*, 25, 509–530.
- Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from <http://www.afhayes.com/public/process2012.pdf>
- Hopkins, N., Reicher, S., Harrison, K., Cassidy, C., Bull, R., & Levine, M. (2007). Helping to improve the group stereotype: On the strategic dimension of prosocial behavior. *Personality and Social Psychology Bulletin*, 33, 776–788.
- Ip, G. W. M., Chiu, C. Y., & Wan, C. (2006). Birds of a feather and birds flocking together: Physical versus behavioral cues may lead to trait- versus goal-based group perception. *Journal of Personality and Social Psychology*, 90, 368–381.
- 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
- 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.
- Lickel, B. (2012). Retribution and revenge. In L. R. Tropp (Ed.) *The Oxford handbook of intergroup conflict* (pp. 89–105). New York, NY US: Oxford University Press.
- Lickel, B., Hamilton, D. L., & Sherman, S. J. (2001). Elements of a lay theory of groups: Types of groups, relational styles, and the perception of group entitativity. *Personality and Social Psychology Review*, 5, 129–140.

COLLECTIVE RESPONSIBILITY

Lickel, B., Miller, N., Stenstrom, D. M., Denson, T. F., & Schmader, T. (2006).

Vicarious retribution: The role of collective blame in intergroup aggression.

Personality and Social Psychology Review, 10, 372-390.

Lickel, B. & Onuki, M. (2015). Generalization processes in collective responsibility and intergroup conflict. In S. J. Stroessner & J. W. Sherman (Eds.) *Social Perception from Individuals to Groups*. New York, NY: Psychology press.

Lickel, B., Schmader, T., & Barquissau, M. (2004). The evocation of moral emotions in intergroup contexts: The distinction between collective guilt and collective shame. In N. R. Branscombe & B. Doosje (Eds.), *Collective guilt: International perspectives*. New York, NY: Cambridge University Press.

Lickel, B., Schmader, T., & Hamilton, D. L. (2003). A case of collective responsibility: Who else was to blame for the Columbine high school shootings? *Personality and Social Psychology Bulletin*, 29, 194–204.

Lickel, B., Schmader, T., & Spanovic, M. (2007). Group-conscious emotions: The implications of others' wrongdoings for identity and relationships. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and Research* (pp. 351-370). New York: NY: Guilford Press.

Lickel, B., Steele, R. R., & Schmader, T. (2011). Group-based shame and guilt: Emerging directions in research. *Social And Personality Psychology Compass*, 5, 153-163.

Marques, J. M., Abrams, D., & Serôdio, R. G. (2001). Being better by being right: Subjective group dynamics and derogation of in-group deviants when generic

COLLECTIVE RESPONSIBILITY

- norms are undermined. *Journal of Personality and Social Psychology*, 81, 436-447.
- Marques, J. M., Yzerbyt, V., & Leyens, J.-Ph. (1988). The 'Black Sheep Effect': Extremity of judgments toward ingroup members as a function of group identification. *European Journal of Social Psychology*, 18, 1-16.
- Muller, D., & Butera, F. (2007). The focusing effect of self-evaluation threat in coaction and social comparison. *Journal of Personality and Social Psychology*, 93, 194-211.
- Schmader, T., & Lickel, B. (2006a). Stigma and shame: Emotional responses to the stereotypic actions of one's ethnic ingroup. In S. Levin & C. van Laar (Eds.), *Stigma and group inequality: Social psychological approaches* (pp. 281-286). Mahwah, NJ: Lawrence Erlbaum.
- Schmader, T., & Lickel, B. (2006b). The approach and avoidance function of guilt and shame emotions: Comparing reactions to self-caused and other-caused wrongdoing. *Motivation and Emotion*, 30, 43-56.
- Shapiro, J. R. (2012). Types of threats: From stereotype threat to stereotype threats. In T. Schmader & M. Inzlicht (Eds.), *Stereotype threat: Theory, process, and application* (pp. 71-88). New York, NY: Oxford University Press.
- Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review*, 11, 107-130.
- Shepherd, L., Spears, R., & Manstead, A. S. R. (2013). 'This will bring shame on our

COLLECTIVE RESPONSIBILITY

- nation': The role of anticipated group-based emotions on collective action. *Journal of Experimental Social Psychology*, 49, 42–57.
- Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, 48, 813-838.
- Smith, E. R. (1993). Social identity and social emotions: Toward new conceptualizations of prejudice. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in group perception* (pp. 297-315). San Diego, CA, US: Academic Press.
- Spears, R. (2010). Group rationale, collective sense: Beyond intergroup bias. *British Journal of Social Psychology*, 49, 1-20.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.
- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33, 1-39.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology* 1, 149–178.
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72, 271-322.
- Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. B. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, 70, 1256–1269.

COLLECTIVE RESPONSIBILITY

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987).

Rediscovering the social group: A self-categorization theory. Oxford, England: Blackwell.

van Leeuwen, E. (2007). Restoring identity through outgroup helping: Beliefs about international aid in response to the December 2004 tsunami. *European Journal of Social Psychology*, 37, 661-671.

van Leeuwen, E., & Täuber, S. (2011). Demonstrating knowledge: The effects of group status on outgroup helping. *Journal of Experimental Social Psychology*, 47, 147-156.

van Leeuwen, E., van Dijk, W., & Kaynak, Ü. (2013). Of saints and sinners: How appeals to collective pride and guilt affect outgroup helping. *Group Processes and Intergroup Relations*, 16, 781-796.

Wildschut, T., Pinter, B., Vevea, J. L., Insko, C. A. & Schopler, J. (2003). Beyond the group mind: a quantitative review of the interindividual-intergroup discontinuity effect. *Psychological Bulletin*, 129, 698-722.

COLLECTIVE RESPONSIBILITY

Figures

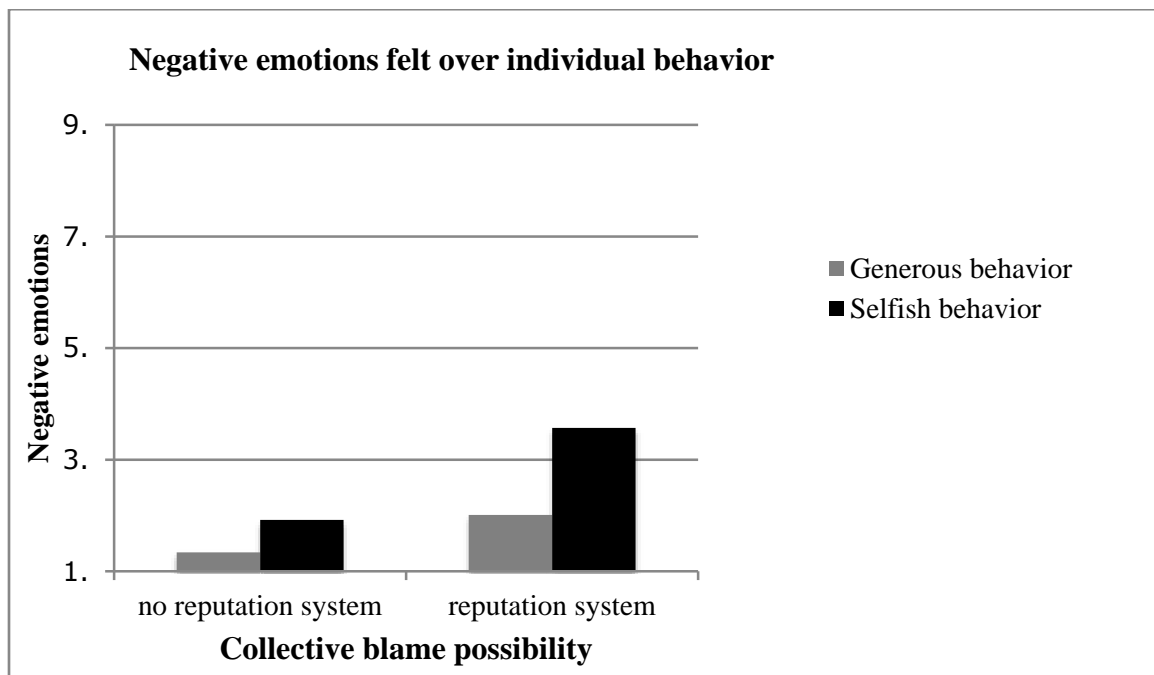


Figure 1: Negative emotions felt over individual behavior as the function of collective blame possibility and the prosocial nature of the behavior.

COLLECTIVE RESPONSIBILITY

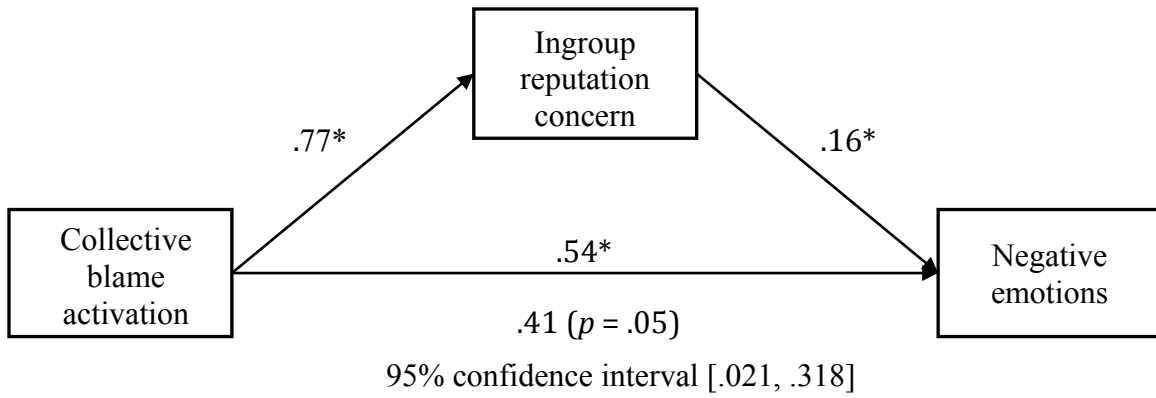


Figure 2: Ingroup reputation concern mediates the effect of collective blame activation on negative emotions. * $p < .05$.

COLLECTIVE RESPONSIBILITY

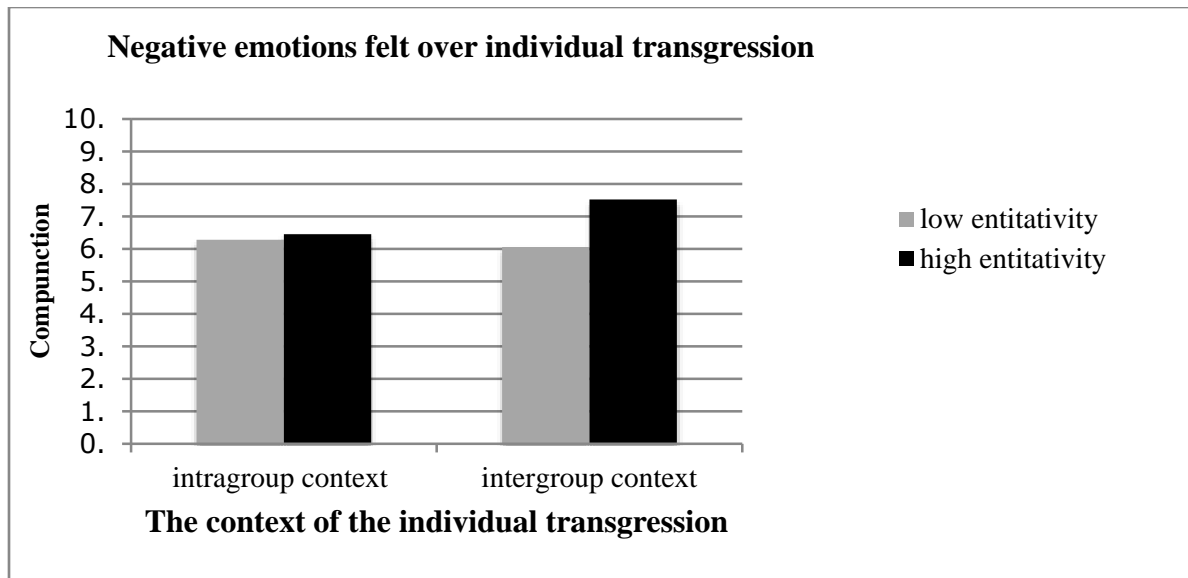


Figure 3: Negative emotions felt over an individual transgression as the function of social context and perceived ingroup entitativity.

COLLECTIVE RESPONSIBILITY

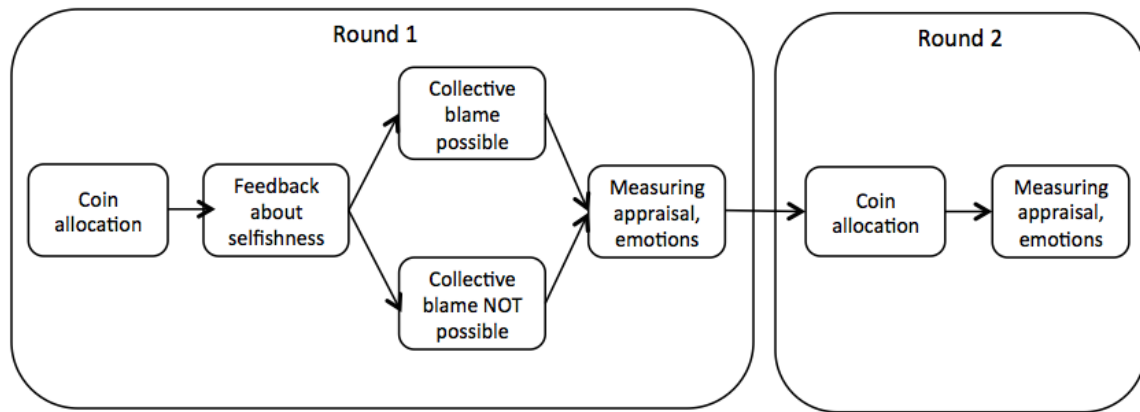


Figure 4: Experiment 4: The timeline of the experimental procedure.

COLLECTIVE RESPONSIBILITY

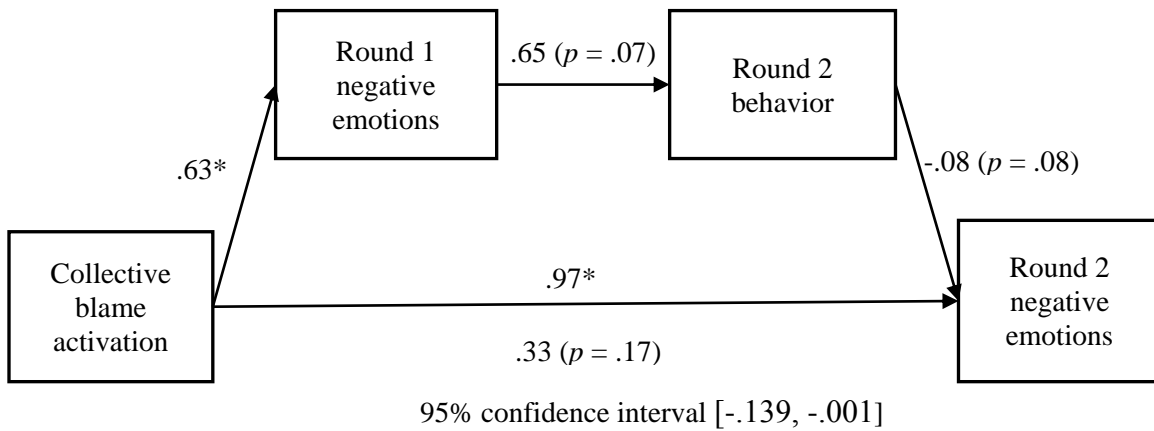


Figure 5: Sequential mediation: Collective blame activation increases negative emotions; negative emotions, in turn, increase compensatory behavior; compensatory behavior reduces subsequent negative emotions. * $p < .05$.