RUNNING HEAD: THIRD-PARTY INTERVENTIONS, ANGER, ATTENTION FOCUS

Costly Third-Party Interventions: The Role of Incidental Anger and Attention Focus in

Punishment of the Perpetrator and Compensation of the Victim

Michaela Gummerum[1]*, Lotte F. Van Dillen[2], Eric Van Dijk[2], Belén López-Pérez[1]

[1] School of Psychology, Cognition Institute, University of Plymouth, UK

[2] Social and Organizational Psychology Unit, Institute for Psychological Research, Leiden University, Netherlands

*Please send correspondence to

Michaela Gummerum

School of Psychology, Cognition Institute

Plymouth University

Drake Circus

Plymouth, Devon, PL4 8AA, UK

Telephone: (+44) (0)1752 584 828

Fax: (+44) (0)1752 584 808

Michaela.gummerum@plymouth.ac.uk

Abstract

Costly third-party interventions have been regarded as hallmarks of moral behaviour, because they are unlikely to be motivated by self-interest. This research investigated the cognitive and emotional processes underlying two types of costly third-party interventions by manipulating incidental emotions and attention focus. In Study 1, we investigated the effect of incidental anger on third-party punishment decisions. Study 2 addressed the effect of incidental anger on third-party compensation decisions. In both studies, participants were induced to either an angry or neutral emotion and then had to wait or were distracted before engaging in third-party interventions. In Study 1, angry participants punished highly unequal distributions significantly more than those in the neutral emotion condition. In Study 2, angry participants compensated highly unfair distributions significantly less than those in a neutral emotion. In both studies, the effect of incidental anger was only significant in the wait, not the distraction condition. Study 3 again focused on third-party compensation decisions. Participants were induced to either a self-focused anger or an other-focused anger emotion, and attention focus (wait, distraction) was manipulated experimentally. Those in the other-focused anger condition compensated significantly more than participants in the self-focused anger condition. These results indicate that (self-focused) incidental anger led to antagonistic responses. Incidental anger was only associated with higher third-party compensation when it includes a focus on a suffering other. For incidental emotions to bias subsequent decisions requires attentional resources.

Keywords: Third-party punishment; third-party compensation; anger; attention

What characterizes a moral person? One central quality of highly moral people is their selflessness or orientation towards others (e.g., Walker & Hennig, 2004). Thus, one trait common to moral exemplars, such as Martin Luther King Jr., Oskar Schindler, or Mother Teresa, is their willingness to act on behalf of others affected by an injustice, harm, or (moral) violation, even when these actions entail significant costs. Such costly third-party interventions have been regarded as a hallmark of moral behavior, because they are unlikely to be motivated by self-interest (Vaish, Missana, & Tomasello, 2011).

Third parties' observations of injustices or moral violations can lead to two types of interventions that help people address the psychological need that "justice has been done" (see Darley & Pittman, 2003; Van Prooijen, 2010): Punishing the violator (i.e., retributive justice) and restoring the harm done by compensating the victim (i.e., compensatory justice). The aim of the present research was to assess the role of incidental anger and attention focus in costly third-party punishment and compensation. This allowed us to investigate the cognitive and emotional processes underlying costly third-party interventions.

Costly third-party punishment and compensation

One experimental procedure developed to measure whether unaffected third parties engage in actual costly punishment towards those who violated the norm of fair sharing is the third-party punishment game (TPP; Fehr & Fischbacher, 2004). In the TPP, Person A first decides whether to allocate money to Person B who can only accept the proposed allocation. After observing A's transfer to B, Person C, the third-party punisher, can decide to punish A by spending some of his/her own endowment. For every monetary unit the punisher spends (e.g., 1 coin), A loses two monetary units (e.g., 2 coins), but the payoff of B is not affected. Thus, punishers spend some of their own endowment to take away resources from A, even though they are not affected by the violation. Punishment is costly as the punishers' returns are lower than if they had not punished. Empirical research with the TPP in diverse societies

(e.g., Fehr & Fischbacher, 2004; Henrich et al., 2006; Nelissen & Zeelenberg, 2009; see Jensen, 2010, for a review) revealed that about 60% of punishers sanctioned A's unequal offers to B. The more unequal the offer by A, the more punishment was administered, suggesting that perceptions of (un)fairness were an important motivator for punishment.

Fewer studies have assessed whether third parties also invest resources to compensate the victims of unfairness. Theoretically, both punishment of violators and compensation of victims might be motivated by the psychological need to "do justice". Whereas punishment is based on a motive for "just deserts" and aimed at having violators pay for what they did, compensating the victims endeavors to restore their situation as closely as possible to the pre-violation state (Darley & Pittman, 2003; Van Prooijen, 2010).

Most studies on third-party compensation gave participants the chance to choose between punishing the violator, compensating the victim, or to engage in both types of interventions. Van Prooijen (2010) reported that third parties generally preferred (hypothetical) punishment over compensation. However, participants who felt emotionally close to the victim were more likely to compensate. Lotz, Okimoto, Schlösser, and Fetchenhauer (2011) found that participants favored compensation over punishment, but the majority engaged in both types of interventions. Leliveld, Van Dijk, and Van Beest (2012) showed that empathic concern towards the victim moderated costly third-party compensation or punishment. Those low in empathic concern were more likely to punish than to compensate, while participants high in empathic concern chose compensation over punishment.

Why engage in costly third-party interventions? The role of (negative) emotions

Why do people engage in costly third-party punishment? In repeated interactions, third parties were more likely to punish, when the victim had the possibility to return the favour and reciprocally engage in third-party punishment in the future (Carpenter &

Matthews, 2012). Third parties who felt an obligation to reciprocate a violator's dishonest behavior increased their hypothetical and actual punishment (Whitson, Wang, See, Baker, & Murnighan, 2015). Yet, in one-shot situations where neither victims nor violators can reciprocate, anger was an important driver of costly third-party punishment, more so than other emotions, such as guilt or self-focused feelings of threat (Fehr & Fischbacher, 2004; Lotz et al., 2011; Nelissen & Zeelenberg, 2009). The relationship between costly third-party punishment and anger is not surprising, given that anger has been defined as an other-condemning moral emotion in response to unjustified insults, unfair treatment and betrayal (Haidt, 2003; Mikula, Scherer, & Athenstaedt, 1998; van den Bos, 2003). Furthermore, anger has been linked to the motivation to take revenge or punish a person who is perceived to have acted unfairly (Darley & Pittman, 2003; Montada & Schneider, 1989).

At first glance, the effect of anger on third-party compensation seems to be less straightforward. Emotion research has typically conceptualized anger as a negative emotion that leads to antagonistic interpersonal behavior and a decrease in prosocial actions towards others (e.g., Roseman, Wiest, & Swartz, 1994). This runs counter to the idea that third-party compensation is based on focusing on, feeling close to, and experiencing empathic concern for the victim of a violation (e.g., Leliveld et al., 2012; Van Prooijen, 2010). Yet, Van Doorn et al. (2014) proposed that anger is experienced as a reaction to violations of moral standards, fairness, or equality. This type of reaction, also called moral outrage (Batson et al., 2007) or indignation (Carpenter & Matthews, 2012), motivates third parties to restore equality either through punishment or compensation. Indeed, Lotz et al. (2011) showed that self-reported feelings of moral outrage positively predicted both third-party punishment and compensation.

Previous research on the effect of anger on third-party interventions has mainly studied anger as an *integral* emotion (Lerner, Li, Valdesolo, & Kassam, 2015) that arises as part of the decision-making situation (e.g., as a reaction to the unfair treatment of others). The

current studies examined the effect of *incidental* anger on third-party interventions. Incidental emotions, triggered in one situation, have been shown to "carry over" and bias behaviors or decisions in other, unrelated situations (see Lerner et al., 2015). Focusing on incidental (rather than integral) anger can not only reveal whether anger influences third-party interventions, but can also shed light on the emotional and cognitive processes underlying third parties' decisions. According to appraisal theories of emotions (e.g., Frijda, 1986; Lerner et al., 2015; Smith & Ellsworth, 1985) different emotions can be distinguished according to a range of cognitive dimensions (or appraisals). The appraisal-tendency hypothesis (e.g., Lerner et al., 2015; Lerner & Keltner, 2000; Tiedens & Linton, 2001) proposes that incidental emotions trigger emotion-specific appraisal dimensions which activate a cognitive predisposition to evaluate future situations or events in line with the emotion's underlying appraisal patterns. A number of studies (e.g., Lerner & Keltner, 2000; Keltner, Ellsworth, & Edwards, 1993; Tiedens & Linton, 2001) investigated the effect that incidental emotions characterized by appraisal patterns have on subsequent judgments and choices. Most pertinent to the current research, incidental anger affected the severity of people's moral judgments (Seidel & Prinz, 2013), led to higher punitive attributions (Lerner, Goldberg, & Tetlock, 1998), and higher second-party punishment when the punisher was the victim of unfairness (Seip, Van Dijk, & Rotteveel, 2014). Thus, the first aim of our research was to assess whether the appraisal dimensions triggered by incidental anger do carry over and activate appraisal-consistent tendencies in third parties' punishment and compensation.

Incidental emotions and attention

The second aim of our studies was to examine under what conditions incidental anger might bias third-party interventions. Research suggests that the subjective experience of emotions requires mental resources (Kron, Schul, Cohen & Hassin, 2010; Van Dillen, Heslenfeld, & Koole, 2009; Van Dillen & Koole, 2007). Cognitive, and particularly

attentional, resources might be especially critical when understanding the effect of incidental emotions on subsequent judgments and behaviors. Van Dillen, Van der Wal and Van den Bos (2012) argued that the incidental emotional response only carries over and biases subsequent cognitions and behaviors when it is sustained through attentional processes. As illustrated in Figure 1, an incidental emotional response (box 1) triggers sustained emotion processing (box 3), and emotion-congruent cognitions and behaviors (box 4), only if the initial emotion has captured attention (box 2). These attentional processes vary based on situational demands or individual dispositions. Van Dillen et al. (2012, Studies 1, 2) found that participants with weak attentional control were less able to disengage their attention from their disgust emotions and, consequently, exhibited harsher moral judgments. In Study 3, Van Dillen et al. (2012) manipulated attention by asking participants after a disgust induction to either focus on how they felt (feel condition – intensified attention to incidental emotion), to wait and watch a blank screen (wait condition – sustained attention to incidental emotion) or to engage in a game of Tetris (distraction condition – disengaging attention from incidental emotion). Participants in the distraction condition reported the mildest moral judgments indicating that attention focus modulated the effect of incidental disgust on subsequent moral cognitions.

To our knowledge, only one study investigated the relationship between attention, third-party interventions, and anger, albeit integral anger. Wang et al. (2011) presented third parties with a trust game, in which a second player did not reciprocate trust by the first player. Participants engaged in more third-party punishment and reported more integral anger when they had to make their decision immediately compared to conditions in which they engaged in a delay and distraction task before making their punishment decisions. It is unclear whether and how these findings extend to incidental anger, however, which will be the focus of the present research.

The present research

The present research aimed to investigate whether and under what conditions incidental anger affects third-party interventions. Studying the role of incidental rather than integral anger for third-party interventions allowed us to examine some of the emotional and cognitive processes underlying third parties' decisions. We suggest that incidental anger carries over and biases third-party interventions, but only when participants have attentional resources (Figure 1; Van Dillen et al., 2012).

Three studies were conducted examining the effect of incidental anger on third-party punishment (TPP; Study 1) and third-party compensation (TPC; Studies 2, 3). In contrast to previous research which gave third parties the opportunity to punish and/or compensate simultaneously (e.g., Lotz et al., 2011), we investigated third-party punishment and compensation separately. This gave us the opportunity to assess the relationship between the appraisal of incidental anger and its effect on subsequent third-party interventions in more detail. Appraisal theories of emotions (e.g., Smith & Ellsworth, 1985) have suggested that incidental anger is appraised as perceiving another agent being responsible for a (negative) outcome and triggers antagonistic responses and blame (e.g., Seidel & Prinz, 2013; Seip et al., 2014; Small & Lerner, 2008). Accordingly, incidental anger should only enhance punishment, but not compensation tendencies. However, Van Doorn et al. (2014) conceptualized anger as a reaction to unjust or unfair situations (i.e., moral outrage, indignation) that triggers the tendency to restore fairness. In that sense, anger is appraised as perceiving the blockage of a desired goal (e.g., achieving equality) and triggers an effort to remove the obstruction (Carver & Harmon-Jones, 2009). Consequently, incidental anger should enhance both punishment and compensation tendencies.

Most of the earlier studies on third-party interventions and (integral) anger asked participants to respond to one scenario in which the violator allocated resources unfairly or engaged in non-reciprocal behavior (e.g., Lotz et al., 2011; Nelissen & Zeelenberg, 2009,

Wang et al., 2011). In Studies 1 and 2, participants had to make a binding punishment (compensation) decision for each possible distribution in the TPP (TPC) game. Using this strategy method (Fehr & Fischbacher, 2004) allowed us to examine whether incidental anger was associated with more heuristic processing and certain obliviousness to situational factors (Lerner et al., 1998; Small & Lerner, 2008). If this is the case, then incidental anger should trigger similar levels of punishment (compensation) across distributions. However, if angry third parties take situational circumstances into account, then the effect of incidental anger should be more pronounced for the punishment (compensation) of unfair than for fair or advantageous distributions.

Study 1: Third-party punishment

Study 1 investigated whether costly third-party punishment of unfair allocations was influenced by incidental anger and distraction. After participants engaged in an emotion induction task aimed at eliciting an angry or neutral incidental emotion, they were asked to either wait (wait condition) or play a game of Tetris (distraction condition; Van Dillen et al., 2012). Thus, in the wait condition, participants had the opportunity to sustain attention to the incidental emotion, whereas playing Tetris in the distraction condition aimed to disengage attention from the incidental emotion. We expected that incidental anger would lead to higher third-party punishment of unfair offers than a neutral emotion in the wait condition. In the distraction condition, we did not expect to find a difference in punishment between participants induced to an angry or a neutral emotion.

Method

Participants[1]

One-hundred and forty-nine adults (109 females, 40 males; $M_{Age} = 24.98$ years, $SD = 10.61$ years, age range $18 - 67$ years) from southern England participated. Participants were

recruited from a participant pool and took part either for course credit or financial rewards and had a chance for receiving extra cash vouchers.

Design

A two (attention focus: wait vs. distraction) by two (emotion condition: neutral vs. anger) between-subjects design was employed. Participants were randomly allocated to one of four conditions: wait/neutral ($n = 40$), wait/anger ($n = 38$), distraction/neutral ($n = 34$), and distraction/anger ($n = 37$). The dependent variables were participants' punishment decisions in the TPP.

Procedure

Up to five participants were tested simultaneously in the laboratories of the authors' university. Participants were seated at computer terminals in separate cubicles and were asked to enter a personal ID code, their date of birth and gender. Once they signed the consent form participants received instructions for the TPP. They were told that in addition to the show-up fee (£4 per 30 minutes or course credit) the points distributed in that game would be converted into raffle tickets with the chance to win one (or more) of 50 £20 amazon vouchers. The more points/raffle tickets participants accrued, the higher their chance to win a voucher. Participants' final points payoffs were determined by matching one randomly selected decision with the decision of an anonymous interaction partner.

After the TPP game instructions, participants had to complete two sets of quiz questions. Each question represented an example distribution of points between Person A and Person B and an example of how many points Person C paid to punish Person A. Participants had to calculate the correct payoffs for each player in the game, depending on how many points Person C paid to take away points from Person A. Incorrect answers received an automatic prompt, and after three such prompts participants received further instructions and completed the quiz again.

Next, participants were presented with the emotion induction manipulation using a recalling event procedure where they had to generate and think about an event in their own lives that elicited a certain emotion. This procedure has been shown to be the most successful technique in eliciting a specific emotion (Brenner, 2000). In the neutral conditions, participants were instructed to remember and write about their dinner the night before the experiment. In the anger conditions, participants were instructed to remember and write about an event that made them feel "furious". Participants had to write at least 50 characters about the events for at least 3 minutes. After that, the screen changed automatically to the attention focus manipulation. Participants in the wait condition watched a blank screen and were instructed to wait until the next task loaded for 3 minutes; participants in the distraction condition played a game of Tetris for 3 minutes (see Van Dillen et al., 2012).

We used the strategy method (e.g. Fehr & Fischbacher, 2004) to record third-party punishment decisions (see Materials, below). Subsequently, participants reported their disgust, anger, sadness, and happiness on 7-point Likert scales (1 = not at all, 7 = very much). Participants then made one decision as Person A.[2] Finally, participants were presented with a clip from the movie The Junglebook to induce a happy mood (von Leupoldt et al., 2007) before being debriefed. After data from all participants was collected, the winning raffle ticket holders were determined and contacted and vouchers were allocated.

Materials[1]

Third-party punishment game. The third-party punishment (TPP) game is a sequential three-player game. At Step 1, Person A decides whether and how to allocate 10 points between him-/herself and an anonymous Person B, who can only accept Person A's allocation choice. At Step 2, Person C, the punisher, is informed about Person A's choice. Person C can then decide whether to pay any points from his/her 5-point endowment to

punish Person A. For every point Person C pays, 2 points are taken away from Person A's payoff. Person B's payoff is not affected by Person C's choice.

Participants were allocated to the role of Person C. To gain a comprehensive picture of Person Cs' behaviors, their choices were assessed using the strategy method (e.g., Fehr & Fischbacher, 2004): Participants were presented with each possible distribution of 10 points between Persons A and B (i.e., Person A gives Person B 0, 1, 2,…10 out of 10 points). For each distribution they had to indicate whether they liked to pay any of their 5 points to take away points from Person A. Thus, participants had to make eleven decisions in eleven different counterbalanced TPP games. These decisions were binding, meaning that participants' choices would be matched and compared to the distribution decision made by Person A. We randomly picked one Person A and noted his/her allocation to Person B. Then it was determined whether and how the Person C (the participant) punished this decision; and C's punishment decision was used as the basis for calculating the final payoff.

Analyses

The linear mixed-effects models procedure was run in SPSS with participants' self-reported emotions and third-party punishment analyzed as continuous variables. Emotion Condition (neutral, anger) and Attention Focus (wait, distraction) were entered as fixed effects. In addition, Emotion Type (disgust, anger, sadness, happiness) was entered for the analyses of the self-reported emotions (i.e. the manipulation check), and Offer (points offered by Person A: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10) was entered for the analyses of third-party punishment. Emotion Type and Offer were also entered as a repeated effect to adjust for potentially correlated data in the respective analyses. For third-party punishment, the model included those predictor variables and interactions that allowed for testing the expected effects. Specifically, the model tested the main effects of Offer and Emotion Condition and the interaction effects of Offer x Emotion Condition, Emotion Condition x Attention Focus,

and Offer x Emotion Condition x Attention Focus. Using a deviance test, the fit of the predicted model was compared to the fit of a full model which included all main and interaction effects of the predictor variables.

<center>Results and Discussion</center>

Participants who failed to answer the two sets of quiz questions correctly were dropped from the analyses. Six participants were dropped from the wait/neutral (final $n = 34$), 1 from the wait/anger (final $n = 37$), 3 from the distraction/neutral (final $n = 31$), and 2 from the distraction/anger conditions (final $n = 35$).

Manipulation check

Table 1 shows the means, standard deviations, and confidence intervals of participants' self-reported emotions. The analyses of the full factorial mixed model (AIC = 1673.37; BIC = 1677.64) revealed that the intensity of participants' emotions varied as a function of Emotion Type, $F(3, 532) = 83.19$, $p < .001$. Participants overall reported more intense happiness than anger, disgust, and sadness ($ps < .001$) whereas they still reported more intense anger than disgust and sadness ($ps < .001$). There was also a main effect of Attention Focus, $F(1, 532) = 8.89$, $p = .003$, with participants reporting more intense emotions in the wait than in the distraction condition. There was a significant interaction of Emotion Type x Emotion Condition, $F(3, 532) = 26.34$, $p < .001$, which was further qualified by a three-way interaction between Emotion Type, Emotion Condition, and Attention Focus, $F(1, 532) = 12.72$, $p < .001$. No other significant main or interaction effects were observed.

Pairwise comparisons were conducted to further analyze the three-way interaction. These revealed that following an anger induction, participants reported significantly more anger and significantly less happiness than in the neutral condition ($ps < .001$), but this was only the case when participants were instructed to wait. When participants' attention had been diverted by a game of Tetris, incidental anger neither significantly increased anger (p

=.287) nor decreased happiness (p =.072) compared to the neutral control condition.

Participants moreover reported significantly more anger than disgust, sadness, or happiness

(ps <.001) following the anger induction, whereas in the neutral conditions, anger reports did

not differ significantly from reports of disgust or sadness (ps > .132) and were significantly

less intense than reports of happiness (ps< .001). Together, these results confirm the

effectiveness of our anger induction.

Third-party punishment

Table 2 shows the parameter estimates and goodness-of-fit statistics for the fixed

effects predicting participants' third-party punishment. Third-party punishment decreased

with increasing offers, $F(1, 1499) = 527.30$, $p < .01$, and angry participants punished more

than participants in the neutral emotion condition, $F(1,1499) = 113.27$, $p < .01$. These main

effects were qualified by statistically significant interactions. The effect of incidental anger

on third-party punishment was more pronounced for low and unfair offers than for fair and

high offers (Offer x Emotion Condition, $F(1, 1499) = 76.81$, $p < .01$). As shown in Figure 2a,

in the wait condition, angry participants punished more than those in the neutral condition.

However, participants' incidental emotion did not affect punishment in the distraction

condition (Figure 2b; Emotion Condition x Attention Focus, $F(1, 1499) = 129.52$, $p < .01$).

The effect of anger on punishment in the wait condition was particularly pronounced for

small and unfair offers (Offer x Emotion Condition x Attention Focus, $F(1, 1499) = 86.86$, p

$< .01$). A deviance test indicated that the fit of the predicted model was significantly worse

than the fit of the full model (Deviance~Full model~ $= 3861.46$), $\chi^2(2) = 7.37$, $p < .05$.

Overall, Study 1 found support for the predicted hypotheses. Participants in the anger

condition punished significantly more than participants in the neutral condition. Thus, anger

elicited in an independent situation carried over and led third parties to engage in more costly

punishment. However, the effects of incidental anger on third-party punishment were not

uniform but depended on the punishers' attentional capacities and situational circumstances. In line with Van Dillen et al. (2012), incidental anger affected punishment in the wait condition, where participants had the attentional resources to continue processing the incidental emotion, but not in the distraction condition, where attention from the incidental emotion was interrupted. Furthermore, in the wait condition, incidental anger particularly affected the punishment of unfair distributions, but not of fair or advantageous distributions. In sum, both people's attentional resources and the fairness of the offer affected the influence of incidental anger on costly third-party punishment.

Study 2: Third-party compensation

While both costly third-party punishment and costly third-party compensation can be regarded as altruistic actions, third-party compensation involves focussing on the situation of and restoring the harm done to the victim of a violation (Darley & Pittman, 2003). Attending to the plight of a victim has traditionally been associated with empathy (e.g., Hoffman, 2000). Indeed, third parties who felt empathic concern or close to the victim engaged in higher compensation behaviour (Leliveld et al., 2012; Van Prooijen, 2010).

It is less clear how anger relates to third-party compensation. Some research on empathy and its relation to altruistic behaviour suggested that individuals who were prone to feel negative emotional arousal in social situations exhibited less empathic concern than those who could regulate their negative emotions (Eisenberg et al., 1994). Seidel and Prinz (2013) found that participants induced to an angry emotion reported a lower sense of obligation to help a needy person than those induced to a happy or neutral emotion. Incidental anger decreased the amount of welfare participants recommended should be given to a hypothetical needy person compared to a neutral emotion (Small & Lerner, 2008). These findings suggest that angry third parties might be less likely to engage in costly compensation than those in a neutral emotion. However, more recently Van Doorn et al. (2014) conceptualized anger as a

reaction to inequality and unfairness, which can lead to both increased third-party punishment and third-party compensation. Similarly, Batson et al. (2007) proposed that unfairness in third-party situations can trigger two types of angry responses: moral outrage as a reaction to violations of fairness standards which should activate the restauration of fairness; or empathic anger triggered by the obstruction of a cared-for person's interests, which should lead to support for this person's benefits. Both moral outrage and empathic anger could lead to increased third-party compensation.

So far, the role of incidental anger for third-party compensation has not been explored. People's appraisals seem to be critical for understanding whether incidental anger carries over and biases third-party compensation. If incidental anger is conceptualized as moral outrage (e.g., Van Doorn et al., 2014) and thus appraised as the blockage of a desired goal (i.e., equality/fairness), it should trigger efforts to remove the obstruction and lead to enhanced third-party compensation. If incidental anger is appraised as perceiving another agent being responsible for (negative) outcomes and triggers antagonistic responses and blame, it should negatively affect third-party compensation compared to a neutral mood.

Study 2 again manipulated attention focus after the emotion induction. Incidental anger should only affect third-party compensation in the wait, but not the distraction condition, because only in the wait condition do participants have the attentional resources to keep on processing the incidental emotion. Moreover, Study 2 again varied the fairness of the offer. Compensation rates should be higher the smaller the allocation to the receiver.

<div align="center">Method</div>

Participants[1]

One-hundred and forty-eight adults (97 females, 51 males; M_{Age} = 25.29 years, SD = 10.10 years, age range 18 – 66 years) from southern England participated for course credit or financial rewards.

Design

The experimental design was identical to Study 1. Participants were randomly allocated to one of the four conditions, wait/neutral ($n = 39$), wait/anger ($n = 38$), distraction/neutral ($n = 36$), and distraction/anger ($n = 35$). The dependent variables were participants' compensation decisions in the third-party compensation game.

Procedure

The procedure was the same as in Study 1 with the exception of participants making 11 decisions in a third-party compensation game rather than a third-party punishment game.

Materials[1]

Third-party compensation game. The third-party compensation (TPC) game is a sequential three-player game. At Step 1, Person A decides whether and how to allocate 10 points between him-/herself and an anonymous Person B, who can only accept. At Step 2, Person C is informed about Person A's choice and can then decide whether to pay any points from his/her 5-point endowment to compensate Person B. For every point Person C pays, 2 points are given to Person B. Person A's payoff is not affected by Person C's choice.

Participants were instructed to the role of Person C and had to make eleven binding decisions in TPC games (presented in counterbalanced order) for each possible outcome (i.e., for the possibilities that Person A offered 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 out of 10 points to Person B, respectively). For each game, Person C's endowment was 5 points.

Analyses

Analysis strategy was the same as in Study 1. For third-party compensation, the model included the predicted main effects of Offer and Emotion Condition and the predicted interaction effects of Offer x Emotion Condition, Emotion Condition x Attention Focus, and Offer x Emotion Condition x Attention Focus.

<div align="center">Results and Discussion</div>

Participants who did not answer the two sets of quiz questions correctly were dropped from the analyses. Five participants were dropped from wait/neutral (final $n = 34$), 3 from the wait/anger (final $n = 35$), 3 from distraction/neutral (final $n = 33$), and none from distraction/anger (final $n = 35$).

Manipulation check

Table 1 shows the means, standard deviations, and confidence intervals for participants' self-reported emotions. The full factorial mixed model analyses (AIC = 1720.65, BIC = 1724.93) revealed that the intensity of participants' emotions varied as a function of emotion condition, $F(3, 532) = 64.71$, $p < .001$. Participants overall reported most intense happiness (ps <.001), but they still reported more intense anger than disgust or sadness (ps <.001). There was a significant interaction between Emotion Type and Emotion Condition, $F(3, 580) = 33.28$, $p < .001$, which was further qualified by a three-way interaction between Emotion Type, Emotion Condition, and Attention Focus, $F(3, 532) = 16.61$, $p < .001$. Pairwise comparisons revealed that following the anger induction, participants reported significantly more anger and significantly less happiness than in the neutral condition (ps <.001), but this was only the case when participants were instructed to wait. In the distraction condition, incidental anger no longer increased angry feelings (p >.302) and only resulted in a marginally significant decrease in happiness ($p = .066$) relative to the neutral control condition. Participants moreover reported significantly more anger than disgust, sadness, or happiness (ps <.001) following the anger induction, whereas in the neutral conditions, anger reports did not differ from reports of disgust or sadness (ps > .123) and were significantly *less* intense than reports of happiness (ps< .001). Together, these results suggest that our incidental anger manipulation was again effective.

Third-party compensation

Parameter estimates for the fixed effects predicting participants' third-party compensation are displayed in Table 2. Third-party compensation decreased with increasing offers, $F(1, 1500) = 797.20$, $p < .01$. Angry participants compensated less than participants experiencing a neutral emotion, $F(1,1500) = 52.93$, $p < .01$, and this effect was more pronounced for low and unfair offers than for fair and high offers (Offer x Emotion Condition), $F(1, 1500) = 23.92$, $p < .01$. As shown in Figure 3a, in the wait condition, angry participants compensated less than participants in the neutral emotion condition. However, there was no difference in the compensation between participants experiencing an angry or neutral incidental emotion in the distraction condition (Figure 3b; Emotion Condition x Attention Focus, $F(1, 1500) = 11.93$, $p = .001$). In the wait condition, the effect of incidental anger was particularly pronounced for low and unequal offers, but not for fair and equal offers. In the distraction condition, there were no differences in compensation for participants induced to an angry or neutral incidental for either low or high offers (see Figure 4; Offer x Emotion Condition x Attention Focus, $F(1, 1500) = 6.44$, $p = .01$). A deviance test showed that the fit of the predicted model was not significantly different to the fit of the full model (Deviance$_{\text{Full model}}$ = 3583.02), $\chi^2(2) = 5.58$, $p > .05$.

Thus, concerning the role of distraction, Study 2 replicated the effect of Study 1 as well as Van Dillen et al.'s (2012) research. In the distraction condition there was no difference in compensation between third parties induced to an angry or neutral emotion. Importantly, playing Tetris reduced participants' anger but it did not enhance participants' positive feelings, which rules out the alternative explanation that playing Tetris functions as a sort of mood repair (see also Van Dillen et al., 2012). Thus, by blocking people's attentional resources, distraction seems to affect the continued processing of incidental emotions and, accordingly, subsequent altruistic third-party compensation behavior, particularly for low and unfair offers (see Figure 1).

Concerning the effect of incidental anger on third-party compensation, we found that angry third parties compensated a victim less than those induced to a neutral emotion. Thus, in the current study incidental anger might have initiated an antagonistic action tendency that runs counter to focusing on, caring for, and compensating the victim. This contradicts previous research on integral anger and third-party punishment and compensation (Lotz et al., 2011; Van Doorn et al., 2014). How can we reconcile these findings? Incidental anger per se might not increase compensation, but third parties might additionally have to appraise the plight of the victim (empathic anger). Vitaglione and Barnett (2003) defined empathic anger as the anger experienced and directed at a transgressor on behalf of a suffering victim, and Batson et al. (2007) suggested that empathic anger most likely occurs when the interests of a person for whom one cares have been thwarted. This can be contrasted with personal anger, that is the "anger one might feel when one's own interests are thwarted" (Batson et al., 2007, p. 1273). In Vitaglione and Barnett's (2003) study empathic anger positively predicted hypothetical helping towards a victim. These results also chime with Leliveld et al.'s (2012) findings that trait empathic concern was related to third-party compensation. The emotion induction procedure employed in Study 2 might not have provoked such empathic or other-focused anger. Indeed, inspections of the narratives participants produced during the anger induction task in Study 2 indicate that while all descriptions referred to the violation of a social, moral, or fairness standard, only a minority of participants (15%) recalled events where something negative happened to another person.

Study 3: Other- versus self-focused anger and third-party compensation

The results of Study 2 suggest that for anger to lead to increased third-party compensation it may need to include an empathic or other-focus (see also Leliveld et al., 2012). So far, very little research has systematically investigated how different types of anger affect third-party compensation, and all of these studies assessed anger as an integral

emotion. Vitaglione and Barnett (2003) found that empathic anger predicted participants' desire to help a victim, but this study did not include any measures of self-focused anger. Lotz et al. (2011) showed that offender-focused moral outrage correlated positively with third-party compensation. Yet, from this study it is difficult to distinguish whether participants might have additionally experienced self- or other-focused anger and whether these emotions might differentially influence third-party compensation.

The main goal of Study 3 was to study the effects of self- and other-focused incidental anger on third-party compensation. Adapting research by Batson et al. (2007), participants in the self-focused anger condition were asked to recall an event that made them feel furious because they had been undeservedly harmed or treated unfairly. Participants in the other-focus anger conditions were asked to write about an event that made them feel furious because another person had been undeservedly harmed or treated unfairly. Such an other-oriented response to the suffering or bad treatment of another has been regarded as the key dimension of empathy (e.g., Lamm, Batson, & Decety, 2007).

Participants could engage in the third-party compensation of the recipient of a highly unfair distribution (recipient gets 0 out of 10 points), an unfair distribution (recipient gets 2 out of 10 points), and a fair distribution (recipient gets 5 out of 10 points). If incidental other-focused anger triggers predisposition to empathically pay attention to others' afflictions, then we would expect more third-party compensation in the other-focused compared to the self-focused anger condition, particularly for highly unfair and unfair distributions. Based on the findings of Studies 1 and 2, we additionally expected that self- and other-focused incidental anger would carry over and differentially affect third-party compensation of highly unfair and unfair offers in the wait, but not the distraction condition.

<div align="center">Method</div>

Participants[1]

One-hundred and forty adults (96 females, 44 males; M_{Age} = 24.98 years, SD = 10.80 years, age range 18 – 69 years) took part in the study for financial rewards or course credit.

Design

The experimental design consisted of a two (attention focus: wait vs. distraction) by two (emotion condition: self-focused anger, other-focused anger) between-subjects design. Thirty-five participants were randomly allocated to one of the four conditions, wait/self-focused anger, wait/other-focused anger, distraction/self-focused anger, and distraction/other-focused anger, respectively. The dependent variables were participants' compensation decisions in the third-party compensation game.

Procedure

The experimental set-up, TPC instructions, and TPC Quiz questions were the same as in Study 2. Before participants made their decisions in the TPC game, they were presented with the emotion induction manipulation. In the self-focused anger condition, participants were asked to recall and write about an event that made them feel furious, because they had been undeservedly harmed or treated unfairly. In the other-focused anger condition, participants were asked to recall and write about an event that made them feel furious because a person for whom they cared had been undeservedly harmed or treated unfairly (see Batson et al., 2007). Participants had to write at least 50 characters for at least 3 minutes. Then the screen changed automatically to the attention focus manipulation, which was identical to the ones employed in Studies 1 and 2.

Next, participants were asked to make three decisions as Person 3 in the TPC. They were presented with three distributions of points between Persons A and B in counterbalanced order: (1) Person A gave 0 out of 10 points to Person B (extremely unequal distribution); (2) Person A gave 2 out of 10 points to Person B (unequal distribution); (3) Person A gave 5 out of 10 points to Person B (equal distribution).   For each game, Person

C's endowment was 5 points, and they were asked to make a binding compensation decision for each of the three distributions.

To assess the success of the experimental manipulations, participants had to subsequently complete the situational empathic concern scale from the Empathic Response Questionnaire (Batson, Fultz, & Schoenrade, 1987) and the offender-focused anger scale (Lotz, et al., 2011). After that, participants completed two items to assess current rumination (López-Pérez & Ambrona, 2015). Participants then made one decision as Person A[2] before being presented with a clip from the movie The Junglebook to induce a happy mood. Finally, participants were thanked and debriefed. After data from all participants was collected, the winning raffle ticket holders were determined and contacted and vouchers were allocated.

Materials[1]

Third-party compensation game. The structure of the third-party compensation (TPC) game was identical to Study 2. Punishers/Persons C had to make binding decisions in three TPC games (i.e., Person A offered 0, 2, or 5 out of 10 points to Person B).

Manipulation check. The 6-item Situational Empathic Concern Scale from the Empathic Response Questionnaire (Batson et al., 1987) evaluates participants' situational empathic concern, that is, to what extent they have felt sympathy and tenderness (i.e., sympathetic, moved, compassionate, tender, warm, softhearted) on a 7-point Likert scale (1= not at all to 7 = extremely; $\alpha = .86$). The 5-item Offender-focused Moral Outrage Scale (Lotz et al., 2011) assesses to what extent participants experience an angry emotional response towards the behavior of another person (i.e., angry, shocked, hostile, distress, aggravated) on a 7-point Liker scale ranging from 1 = not at all to 7 = extremely ($\alpha = .87$). The manipulation check additionally included filler items. Participants were asked to what extent they felt alarmed, grieved, worried, and troubled on the same 7-point Likert scale.

Situational rumination (López-Pérez & Ambrona, 2015). Participants indicated to what extent *they were having negative thoughts* and *they were thinking about their feelings* on a 7-point Likert scale ranging from 1 = not at all to 7 = extremely (α = .70).

Analyses

Analysis strategy for third-party compensation was the same as in Studies 1 and 2. Fixed-effects predictors included in the models were Emotion Condition (self-focused anger, other-focused anger), Attention Focus (wait, distraction), and Offer (points offered by Person A: 0, 2, 5). Offer was also entered as a repeated effect. The model included the predicted main and interaction effects of Offer and Emotion Condition, Offer x Emotion Condition, Emotion Condition x Attention Focus, and Offer x Emotion Condition x Attention Focus.

<div align="center">Results and Discussion</div>

One participant from the distraction/self-focused anger condition was removed from the analyses, because s/he answered both sets of quiz questions incorrectly.

Manipulation check

Concerning empathic concern, a univariate ANOVA with the between-subjects factors Emotion Condition and Attention Focus revealed a significant main effect of Emotion Condition, $F(1, 138) = 4.30$, $p = .04$, $\eta_p^2 = .03$. As shown in Table 3, participants in the other-focused anger condition reported higher empathic concern than those in the self-focused anger condition. The main effect of Attention Focus and the interaction between Emotion Condition x Attention Focus did not reach statistical significance.

Concerning offender-focused moral outrage, a univariate ANOVA revealed a marginally significant main effect of Attention Focus, $F(1, 138) = 2.97$, $p = .09$, $\eta_p^2 = .02$. There was a tendency for participants in the distraction condition to report lower moral outrage (Table 3). The effect of Emotion Condition and the interaction did not reach statistical significance. Participants in the wait and distraction conditions reported similar

levels of rumination (wait: $M = 3.29$, $SD = 1.45$; distraction: $M = 3.51$, $SD = 1.39$, t(137) = .95, $p = .35$).

Third-party compensation

Table 4 shows the parameter estimates and goodness-of-fit statistics for the fixed effects predicting participants' third-party compensation. Participants compensated highly unfair and unfair distributions more than fair ones, $F(1, 409) = 91.51$, $p < .01$ (see Figure 4). Participants in the other-focused anger condition compensated more than participants in the self-focused anger condition, F(1,409) = 11.52, p = .001, and this effect was more pronounced for unequal offers of 0 and 2 points than for fair offers (Offer x Emotion Condition), $F(1, 409) = 6.67$, $p = .01$. The effect of Emotion Condition x Attention Focus was marginally significant, $F(1, 409) = 3.07$, $p = .08$, and the pattern of results paralleled those of Studies 1 and 2. In the wait condition, participants in the other-focused anger condition tended to compensate more than participants in the self-focused anger condition (Figure 4a), but there was no difference between participants in the self- and other-focused anger conditions in the distraction condition (Figure 4b). A deviance test showed that the fit of the predicted model was not significantly different to the fit of the full model (Deviance$_{\text{Full model}}$ = 1196.81), $\chi^2(2) = 3.56$, $p > .10$.

The findings of Study 3 indicate that for incidental anger to increase third-party compensation it needs an empathic or other-oriented component. Participants induced to other-focused anger compensated highly unfair and unfair distributions more than those induced to self-focused anger. No differences in third-party compensation emerged when Person A allocated resources fairly. To our knowledge, this is the first study that assessed the role of different types of anger for third-party compensation. Our findings are in line with research investigating the relationship between third-party compensation and state and

integral empathic anger (Vitaglione & Barnett, 2003) and empathic concern (Leliveld et al., 2012). They can also help reconcile the findings of Study 2 with those of Lotz et al. (2011).

In contrast to Studies 1 and 2, attention focus only marginally affected the relationship between incidental emotions and third-party compensation. However, the results point into the right direction: Participants in the wait condition reported (marginally) higher incidental emotions; those induced to other-focused anger compensated (marginally) more than those induced to self-focused anger in the wait, but not the distraction condition. In sum, the findings across Studies 1-3 suggest that attentional resources might be necessary for the incidental emotion to affect subsequent behavior (Van Dillen et al., 2012).

General Discussion

This research aimed to assess whether and under what conditions incidental anger affects third parties' punishment of the violator of the norm of fair sharing (Study 1) or the compensation of a victim of unfair allocations (Studies 2, 3). Such costly interventions on behalf of another have been regarded as altruistic and moral behavior (Vaish et al., 2011). To our knowledge, no previous research has investigated whether different types of incidental anger affect third-party interventions and how attentional resources and situational circumstances modulate the relationship between incidental anger and third-party punishment and compensation.

Incidental anger and third-party interventions

The appraisal-tendency framework (e.g., Lerner et al., 2015; Lerner & Keltner, 2000) suggests that studying incidental (rather than integral) anger can illuminate the cognitive and emotional processes underlying third-party interventions. According to this framework, the appraisal dimensions triggered by incidental anger should lead third parties to evaluate and act upon future situations in line with the incidental emotion's appraisal dimension. The results of Studies 1 to 3 indicate that incidental anger (at least as manipulated here) was

appraised as another person being responsible for and engaging in an action leading to a negative outcome and activated antagonistic responses (Smith & Ellsworth, 1985). Incidental anger was associated with higher punishment of unfair distributions (Study 1), but only led to higher compensation for the victim, when the anger induction additionally included a focus on a suffering other (Studies 2, 3). These findings are in line with research investigating the relationship between integral anger and third-party punishment (e.g., Fehr & Fischbacher, 2004; Lotz et al., 2011), incidental anger and second-party punishment (Seip et al., 2014), the effect of incidental anger on welfare judgments and moral judgments of need (Seidel & Prinz, 2013; Small & Lerner, 2008), and the influence of trait empathic concern on third-party compensation (Leliveld et al., 2012).

The results of Studies 2 and 3 are inconsistent with Lotz et al.'s (2011) research showing a positive relationship between integral moral outrage and third-party compensation. It is possible that their measure of offender-focused moral outrage might also have assessed other-focused anger, even though this cannot be distinguished in their data. Future replications of Lotz et al.'s (2011) work could include an additional measure of state empathic concern or empathic anger (see Vitaglione & Barnett, 2013) to investigate this question. Furthermore, our findings on incidental anger and third-party compensation do not fit with Van Doorn's et al.'s (2014) suggestion that anger/moral outrage can be appraised as the blockage of the desire to achieve equality which should increase both third-party punishment and compensation. Batson et al. (2007) showed that participants only reported anger when another violated a fairness standard *and* when this violation either thwarted one's own interests or those of a person for whom one cared for. The violation of the fairness standard itself (an appraisal consistent with Van Doorn et al.'s conceptualization of anger) did not produce anger/moral outrage. Batson et al. (2007) propose that moral outrage (i.e., anger at the violation of a standard) might be more likely in moral domains which deal with

violations of (sexual) taboos and propriety standards. Thus, future research might want to investigate whether and how moral outrage affects third-party compensation of the victim of such propriety violations compared to violations of fairness.

The role of distraction and situational demands

It is important to note that the effects of incidental anger on third-party interventions were only observed in the wait condition (albeit marginally for Study 3), where participants had the opportunity to dwell on their induced emotion. In a distraction condition, where participants engaged in a game of Tetris and therefore could divert their attention from the induced emotion, no significant emotion differences emerged. This finding is in line with Van Dillen et al.'s (2012) research which showed that participants in a disgusted mood reported harsher moral judgments when they were in a wait condition or had to ruminate about their feelings (feel condition) than participants in a distraction condition (see also Wang et al., 2011). Overall, the results of the current studies support the idea that the effect of induced emotion requires attentional resources. Cognitive load or distraction seems to reduce people's sustained processing of incidental emotions and, consequently, also affects the impact of incidental emotions on subsequent behaviors and judgments.

Following Van Dillen et al. (2012), it would be interesting to add a "feel condition" in future experiments, where participants are invited to focus on and ruminate on the angry (or neutral) experience they just reported during the emotion induction. We would then expect an even stronger effect of incidental emotions on third-party interventions in the feel compared to the wait or distraction condition (see Van Dillen et al., 2012). Thus, when participants devote cognitive resources to their incidental emotions (e.g., through rumination), these emotions have a stronger impact on subsequent unrelated behaviour. In line with this, work by Bushman and colleagues (2005) has shown how anger rumination facilitates displaced aggression (i.e., aggression at a target unrelated to the initial trigger).

While our research lends support to the appraisal-tendency hypothesis that emotion-specific appraisals carry over and bias subsequent behavior, we did not find that incidental anger was associated with ignoring situational factors, such as the (in-)equality of the distribution between Persons A and B (see Lerner et al., 2015). In all three studies incidental anger did not affect third-party interventions "across the board", but the influence of incidental anger was stronger for unfair offers. The more equal the distribution the less punishment or compensation was administered, in line with previous research (e.g., Fehr & Fischbacher, 2004). This points to the fact that people are rather sensitive to the violation of (distributional) equality or fairness and react to others' unfair treatment even when these interventions are costly for them. Such concerns for equality are not limited to adults but can already be observed when children allocate resources among or punish third parties (e.g., Gummerum & Chu, 2014; Shaw & Olson, 2012). Thus, fairness considerations provide the backbone for third parties' reactions to unequal allocations.

Implications and limitations

Costly third-party interventions, and instances of costly altruistic actions in general, have challenged classical economic theories as these actions contradict the assumptions of people being only motivated to maximize their (selfish) preferences (e.g., Eckel & Gintis, 2010). Our research complements previous investigations that attempted to unearth the proximate causes of human altruism by focusing on the role of social preferences, personality factors, or emotional processes, among others (e.g., Hammerstein, 2003). Our studies indicate that the way third-parties appraise the situation (triggered by an incidental emotion) affects their punishment and compensation.

Appraisal tendencies of incidental emotions do not just bias third-party interventions, but have also been shown to affect second-party punishment (e.g., Moretti & di Pellegrino, 2010). Seip et al. (2014) found that incidental anger increased punishment of a violator by a

victim in the ultimatum game. It is an open question whether the effects of incidental anger are similar in second- and third-party punishment contexts. According to the appraisal-tendency framework, incidental anger should trigger the same appraisals in both contexts and therefore should affect second- and third-party punishment similarly. Conversely, being the victim of a fairness violation might additionally activate integral anger and therefore lead to higher punishment in second- compared to third-party contexts.

Anger, its biological, evolutionary, and neurological underpinnings, its cognitive and motivational components, and associated action tendencies have received much attention (see Carver & Harmon-Jones, 2009; Lerner et al., 2015; Lerner & Tiedens, 2006, Sell, Tooby, & Cosmides, 2009). As discussed above, appraisal theories of emotion (e.g., Smith & Ellsworth, 1985) and particularly the appraisal-tendency hypothesis (e.g., Lerner & Keltner, 2000) can explain the paradoxical effects of incidental anger on third-party punishment and compensation well. The current studies contribute to this research in two important ways. First, many studies in the appraisal-tendency framework have assessed the effect of incidental emotions on (hypothetical) judgments and decisions (e.g., judgments of causality, perceptions of risk). Our studies show that the appraisals triggered by incidental emotions do not only affect such low-cost decisions, but also costly behaviors that have real consequences for decision-makers and recipients. Second, we have identified one possible limit to the influence of incidental emotions on subsequent behavior, namely attentional resources in the decision-maker (see Figure 1).

Studying the role of incidental anger and attention focus on third-party intervention does not only have theoretical but also applied relevance. In legal contexts, third-parties, such as judges or juries, regularly have to make decisions about punishing a perpetrator or compensating the victim of a crime. They are also frequently presented with emotionally arousing materials (e.g., witness or victim statements, crime scene photographs) that have the

potential to carry over and bias subsequent judgments. While relatively few studies have addressed this question (e.g., Salerno & Bottoms, 2009), the current research has the potential to inform the legal system in understanding when and how incidental anger affects third-party punishment and compensation, and how to counter the effects of incidental anger (e.g., by modulating third parties' attention focus; Van Dillen et al., 2012).

Our research is not without limitations. First, Studies 1 and 2 used the strategy method (Fehr & Fischbacher, 2004); participants had to make binding punishment or compensation decisions for the full range of distributions between Persons A and B. While this allowed us to investigate whether incidental anger affected third-party punishment and compensation depending on situational factors, it might have reduced the visceralness of the decision for participants (Brandts & Charness, 2003). It should be noted that research in the appraisal-tendency framework considers the appraisals of emotions as one of the lingering components of the emotion experience (e.g., Keltner et al., 1993). Indeed, many studies in this paradigm asked participants to make multiple judgments or choices after emotion induction. Nevertheless, future research might want to investigate whether the results of our studies hold when third parties have to make one compared to multiple decisions.

The studies presented here focused on situational influences on third-party interventions. However, research indicates that personality factors, such as justice sensitivity, trait empathic concern, perspective-taking, or people's values contribute to altruistic and moral behavior more generally (Edele, Dziobek, & Keller, 2013; Fetchenhauer & Huang, 2004). Future research might investigate how state and trait variables interact and impact on costly third-party interventions to gain a full picture of the proximate causes of human moral behavior.

References

Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Adults' emotional reactions to the distress of others. In N. Eisenberg, & J. Strayer (Eds.), *Empathy and its development* (pp. 163-184). Cambridge, UK: Cambridge University Press.

Batson, C. D., Kennedy, C. L., Nord, L.-A., Stocks, E. L., Fleming, D. A., Marzette, C. M., Lishner, D. A., Hayes, R. E., Kolchinsky, L. M., & Zerger, T. (2007). Anger at unfairness: Is it moral outrage? *European Journal of Social Psychology, 37,* 1272-1285.

Brandts, J., & Charness, G. (2003). Truth or consequences: An experiment. *Management Science*, *49*(1), 116-130.

Brenner, E. (2000). Mood induction in children: Methodological issues and clinical implications. *Review of General Psychology, 4,* 264-283

Bushman, B. J., Bonacci, A. M., Pedersen, W. C., Vasquez, E. A., & Miller, N. (2005). Chewing on it can chew you up: Effects of rumination on triggered displaced aggression. *Journal of Personality and Social Psychology, 88,* 969–983.

Carpenter, J. P., & Matthews, P. H. (2012). Norm enforcement: anger, indignation, or reciprocity?. *Journal of the European Economic Association*, *10*, 555-572.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect. Evidence and implications. *Psychological Bulletin, 135,* 183-204.

Darley, J. M., & Pittman, T. S. (2003). The psychology of compensatory and retributive justice. *Personality and Social Psychology Review, 7*, 324–336.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology, 111,* 225-236.

Eckel, C., & Gintis, H. (2010). Blaming the messenger: Notes on the current state of experimental economics. *Journal of Economic Behavior & Organization*, *73*, 109-

119.

Edele, A., Dziobek, I., & Keller, M. (2013). Explaining altruistic sharing in the dictator

    game: The role of affective empathy, cognitive empathy, and justice sensitivity.

    Learning and Individual Differences, *24*, 96-102.

Eisenberg, N., Fabes, R.A., Murphy, B., Karbon, M., Maszk, P., Smith, M., et al. (1994). The

    relations of emotionality and regulation to dispositional and situational empathy-related

    responding. *Journal of Personality and Social Psychology, 66*, 776–797.

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.

Fehr, E. & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and*

    Human Behavior, 25, 63–87.

Fetchenhauer, D., & Huang, X. (2004). Justice sensitivity and distributive decisions in

    experimental games. *Personality and Individual Differences*, *36*, 1015-1029.

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Gummerum, M. & Chu, M.T. (2014). Outcomes and intentions in children's, adolescents',

    and adults' second- and third-party punishment behavior. *Cognition, 133*, 97-103.

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.

Hammerstein, P. (2003). *The genetic and cultural evolution of cooperation*. Cambridge, MA:

    MIT Press.

Henrich, J., McElreath, R., Ensminger, J., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J.

    C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., &

    Ziker, J. (2006). Costly punishment across human societies. *Science, 312*, 1767-1770.

Hoffman, M.L. (2000). *Empathy and moral development*: *Implications for caring and justice*.

New York: Cambridge University Press.

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B*, *365*, 2635-2650.

Keltner, D., Ellsworth, P. C., & Edwards, K. (1993). Beyond simple pessimism: Effects of sadness and anger on social perception. *Journal of Personality and Social Psychology, 64,* 740-752.

Kron, A., Schul, Y., Cohen, A., & Hassin, R.R. (2010) Feelings don't come easy: Studies on the effortful nature of feelings. *Journal of Experimental Psychology: General, 139*, 520-534.

Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, *19*, 42-58.

Leliveld, M.C., Van Beest, I., & Van Dijk, E. (2012). Punishing and compensating others at your own expense: The role of empathic concern on reactions to distributive injustice. *European Journal of Social Psychology, 42,* 135-140.

Lerner, J. S., Goldberg, J. H., & Tetlock, P. E. (1998). Sober second thought: The effects of accountability, anger, and authoritarianism on attributions of responsibility. *Personality and Social Psychology Bulletin, 24,* 563-574.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgment and choice. *Cognition and Emotion, 14*, 473-493.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology, 66,* 799-823.

Lerner, J. S., & Tiedens, L. Z. (2006). Portrait of the angry decision maker: How appraisal tendencies shape anger's influence on cognition. *Journal of Behavioral Decision Making*, *19*, 115-137.

López-Pérez, B., & Ambrona, T. (2015). The role of cognitive emotion regulation on the vicarious emotional response. *Motivation and Emotion*, *39*, 299-308.

Lotz, S., Okimoto, T. G., Schlösser, T. & Fetchenhauer, D. (2011). Punitive versus compensatory reactions to injustice: Emotional antecedents to third-party interventions. *Journal of Experimental Social Psychology, 47*, 477-480.

Mikula, G., Scherer, K. R., & Athenstaedt, U. (1998). The role of injustice in the elicitation of differential emotional reactions. *Personality and Social Psychology Bulletin, 24*, 769-783.

Montada, L. & Schneider, A. (1989). Justice and emotional reactions to the disadvantaged. *Social Justice Research, 3*, 313–344.

Moretti, L., & di Pellegrino, G. (2010). Disgust selectively modulates reciprocal fairness in economic interactions. *Emotion, 10,* 169-180.

Nelissen, R. M. A. & Zeelenberg, M. (2009). Moral emotions as determinants of third-party punishment: Anger, guilt, and the functions of altruistic sanctions. *Judgment and Decision-Making*, *4*, 543–553.

Roseman, I. J., Wiest, C., & Swartz, T. S. (1994). Phenomenology, behaviors, and goals differentiate discrete emotions. *Journal of Personality and Social Psychology, 67*, 206-221.

Salerno, J. M., & Bottoms, J. L. (2009). Emotional evidence and jurors' judgments: The promise of neuroscience for informing psychology and law. *Behavioral Sciences and the Law, 27,* 273-296.

Schmitt, M., Gollwitzer, M., Maes, J., & Arbach, D. (2005). Justice sensitivity. *European Journal of Psychological Assessment*, *21*, 202-211.

Seidel, A., & Prinz, J. (2013). Mad and glad: Musically induced emotions have divergent impact on morals. *Motivation and Emotion, 37,* 629-637.

Seip, E. C., Van Dijk, W. W., & Rotteveel, M. (2014). Anger motivates costly punishment of unfair behavior. *Motivation and Emotion*, 1-11.

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences, 106,* 15073-15078.

Shaw, A., & Olson, K.R. (2012). Children discard a resource to avoid inequality. *Journal of Experimental Psychology: General, 141,* 382-395.

Small, D. A., & Lerner, J. S. (2008). Emotional policy: Personal sadness and anger shape judgments about a welfare case. *Political Psychology*, *29*, 149-168.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48,* 813-838.

Tiedens, L. Z., & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology, 81,* 973-988.

Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-party moral transgressions. *British Journal of Developmental Psychology, 29*, 124-130.

Van den Bos, K. (2003). On the subjective quality of social justice: The role of affect as information in the psychology of justice judgments. *Journal of Personality and Social Psychology, 85*, 482-498.

Van Dillen, L.F., Heslenfeld D. J., & Koole, S.L. (2009). Tuning down the emotional brain. *NeuroImage, 45,* 1212–1219.

Van Dillen, L. F., & Koole, S. L. (2007). Clearing the mind: A working memory model of distraction from negative mood. *Emotion, 7,* 715-723.

Van Dillen, L.F., van der Wal, R.C., & van den Bos, K. (2012). On the role of attention and emotion in morality: Attentional control modulates unrelated disgust in moral judgements, *Personality and Social Psychology Bulletin, 38*, 1221-1230.

Van Doorn, J., Zeelenberg, M., & Breugelmans, S.M. (2014). Anger and prosocial behavior. *Emotion Review, 6*, 266-273.

Van Prooijen, J.W. (2010). Retributive versus compensatory justice: Observers' preference for punishing in response to criminal offenses. *European Journal of Social Psychology, 40,* 72-85.

Vitaglione, G. D., & Barnett, M. A. (2003). Assessing a new dimension of empathy: Empathic anger as a predictor of helping and punishing desires. *Motivation and Emotion*, *27*, 301–325.

von Leupoldt, A., Rohde, J., Beregova, A., Thordsen-Sörensen, I., Nieden, J.Z., Dahme, B. (2007). Films for eliciting emotional states in children. *Behavior Research Methods, 39*, 606–609.

Walker, L. J., & Hennig, K. H. (2004). Differing conceptions of moral exemplarity: Just, brave, and caring. *Journal of Personality and Social Psychology*, *86*, 629-647.

Wang, C.S., Sivanathan, N., Narayanan, J., Ganegoda, D.B., Bauer, M., Bodenhausen, G.V., & Murnighan, K. (2011). Retribution and emotional regulation: The effects of time delay in angry economic interactions. *Organizational Behavior and Human Decision Processes, 116*, 46-54.

Whitson, J. A., Wang, C. S., See, Y. H. M., Baker, W. E., & Murnighan, J. K. (2015). How, when, and why recipients and observers reward good deeds and punish bad deeds. *Organizational Behavior and Human Decision Processes*, *128*, 84-95.

Footnotes

[1]The sample sizes for Studies 1-3 were determined a priori by power analysis using the program Gpower (Faul, Erdfelder, Lang, & Buchner, 2007). Previous research found that the effect of negative emotions on punishment and/or compensation ranged between $d = 0.75$ and $d = 89$. Based on the power analysis calculations, we planned to test 35 participants per experimental condition. Participants were randomly allocated to one of the four experimental conditions. Testing continued until approximately 35 participants were tested per condition. To be included in the final analyses, participants had to answer at least one out of the two sets of quiz questions correctly. For Study 1, 12 participants were excluded; for Study 2, 11 participants were excluded; for Study 3, one participant was excluded.

One variable was omitted from the analyses and reporting of results in Studies 1 and 2. In both studies we additionally collected participants' attentional control abilities by administering the Attentional Control Scale (Derryberry & Reed, 2002). While this is an established measure that has been used in previous research, the reliability coefficients (Cronbach's α) for this scale were rather low, particularly in Study 1 ($\alpha = .61$ for the Focusing subscale, $\alpha = .53$ for the Shifting subscale). Because of this, the variable was dropped from the analyses. One variable, justice sensitivity (Schmitt, Gollwitzer, Maes, & Arbach, 2005), was excluded from Study 3. Including it did not change the pattern of results (see Supplementary Materials). All other conditions and measures are reported.

[2] Participants' decisions as Person A were only collected to match them to the decision of one randomly selected Person C to determine the Person C's final output (number of points). Since they are not the focus of the current paper, decisions of Persons A are not analysed and reported here.

Table 1.

Emotion induction manipulation check. Means (SDs) and 95% confidence intervals of self-reported disgust, anger, sadness, and happiness on 7-point Likert scales (1 = not at all, 7 = very much) as a function of Attention Focus (Wait, Distraction) and Emotion Condition (Anger, Neutral).

| Subjective feeling | Attention Focus: Wait | | | | Attention Focus: Distraction | | | |
| | Anger | | Neutral | | Anger | | Neutral | |
	M (*SD*)	95% *CI*	*M* (*SD*)	95% *CI*	*M* (*SD*)	95% *CI*	*M* (*SD*)	95% *CI*
Study 1: Third-party punishment								
Disgust	1.76 (1.21)	[1.41, 2.15]	1.29 (.80)	[1.07, 1.61]	1.54 (.95)	[1.25, 1.90]	1.55 (.73)	[1.32, 1.81]
Anger	3.92 (1.19)	[3.56, 4.28]	1.65 (1.10)	[1.30, 2.05]	2.26 (.82)	[1.97, 2.55]	1.97 (.84)	[1.69, 2.67]
Sadness	1.86 (1.29)	[1.48, 2.28]	1.76 (1.26)	[1.39, 2.23	1.54 (.82)	[1.29, 1.82]	1.61 (.76)	[1.34, 1.89]
Happiness	2.60 (1.45)	[2.26, 3.18]	4.38 (1.44)	[3.91, 4.87]	3.06 (1.28)	[2.62, 3.84]	3.55 (1.18)	[3.15, 4.00]
Study 2: Third-party compensation								
Disgust	1.74 (1.01)	[1.41, 2.07]	1.46 (.92)	[1.18, 1.81]	1.80 (1.26)	[1.43, 2.27]	1.51 (.91)	[1.24, 1.83]
Anger	3.83 (1.40)	[3.37, 4.32]	1.31 (.53)	[1.13, 1.48]	2.23 (.88)	[1.94, 2.52]	1.94 (1.25)	[1.52, 2.38]
Sadness	1.97 (1.20)	[1.60, 2.41]	1.54 (.85)	[1.27, 1.84]	1.97 (1.22)	[1.58, 2.42]	1.85 (1.15)	[1.48, 2.24]

| Happiness | 2.46 (1.34) | [2.05, 2.89] | 4.66 (1.21) | [4.22, 5.05] | 3.00 (1.35) | [2.56, 3.46] | 3.51 (1.52) | [3.00, 4.03] |

Table 2.

Estimates (standard errors) and goodness-of-fit statistics of linear mixed effects models predicting participants' third-party punishment (Study 1) and third-party compensation (Study 2)

	Study 1: Third-party punishment	Study 2: Third-party compensation
Intercept	2.22 (.07)**	2.50 (.06)**
Offer: Increasing	-.23 (.01)**	-.26 (.01)**
Emotion condition: Anger	1.22 (.12)**	-.77 (.11)**
Offer x Emotion condition	-.15 (.02)**	.08 (.02)**
Emotion condition x Attention focus	-1.50 (.13)**	.42 (.12)*
Offer x Emotion condition x Attention focus	.18 (.02)**	-.05 (.02)*
AIC	3856.10	3579.45
BIC	3862.41	3584.76
Deviance	3854.10	3577.44

*$p < .05$, **$p < .01$

Table 3.

Study 3: Emotion induction manipulation check. Means (SDs) and 95% confidence intervals of self-reported moral outrage and empathic concern on 7-point Likert scales (1 = not at all, 7 = extremely) as a function of Attention Focus (Wait, Distraction) and Emotion Condition (Self-focused Anger, Other-focused Anger).

	Self-focused anger		Other-focused anger	
	M (SD)	95% *CI*	*M (SD)*	95% *CI*
Attention focus: Wait				
Moral outrage	2.13 (1.27)	[1.74, 2.57]	2.21 (1.23)	[1.84, 2.62]
Empathic concern	2.47 (.90)	[2.17, 2.76]	3.07 (1.25)	[2.64, 3.50]
Attention focus: Distraction				
Moral outrage	1.85 (1.06)	[1.55, 2.23]	1.81 (1.02)	[1.48, 2.17]
Empathic concern	2.59 (.90)	[2.28, 2.87]	2.79 (1.44)	[2.34, 3.27]

Table 4.

Study 3: Estimates (standard errors) and goodness-of-fit statistics of linear mixed effects

models predicting third-party compensation

	Study 3: Third-party compensation
Intercept	1.86 (.18)**
Offer: Increasing	-.82 (.09)**
Emotion condition: Other-focused anger	1.08 (.32)**
Offer x Emotion condition	-.38 (.15)*
Emotion condition x Attention focus	-.64 (.37)†
Offer x Emotion condition x Attention focus	.20 (.17)
AIC	1195.25
BIC	1199.27
Deviance	1193.25

$^{\dagger}p < .10$, *$p < .05$, **$p < .01$
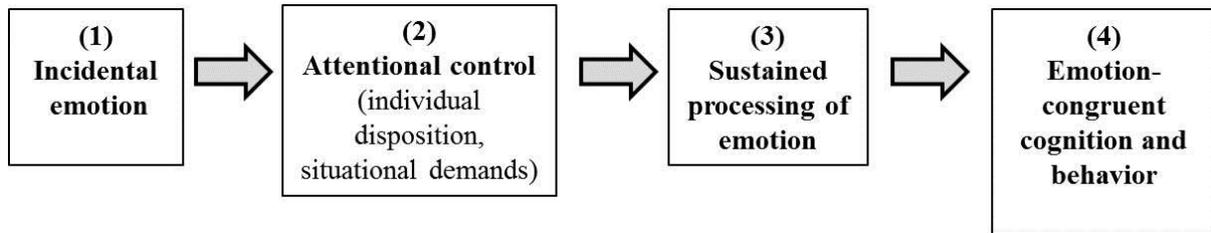
Figure 1. Model on how attentional control impacts the link between incidental emotions and behavior.
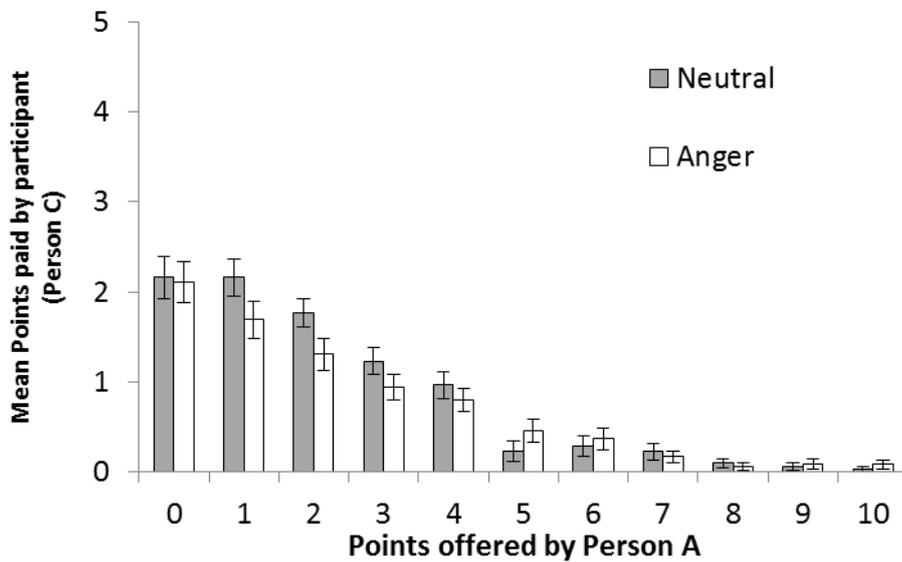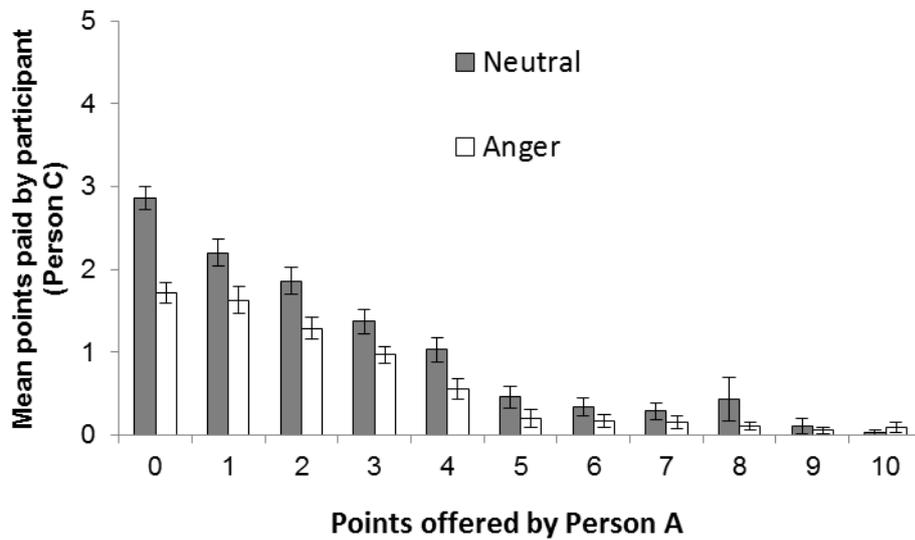
(a)
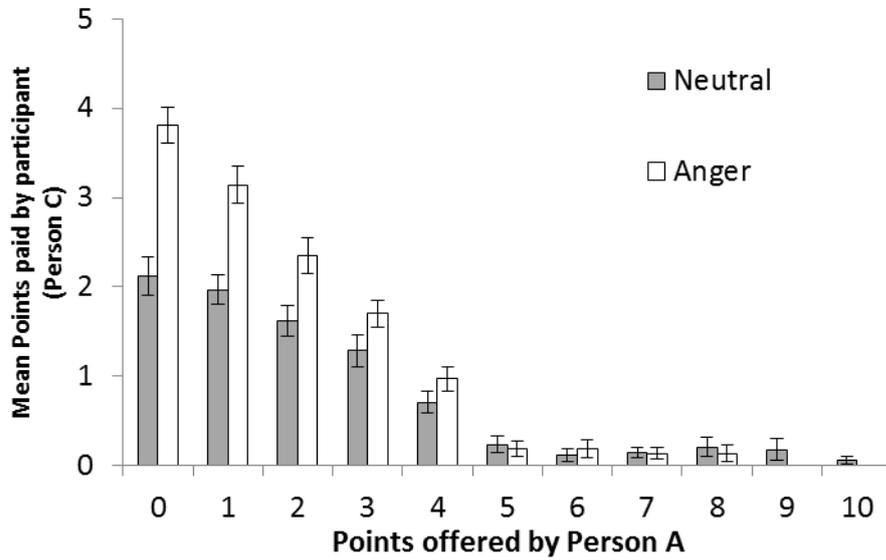
(b)

Figure 2. Study 1: Mean punishment rates (points paid) by Person C in an angry and neutral mood in (a) the wait condition and (b) the distraction condition. Error bars display standard errors.
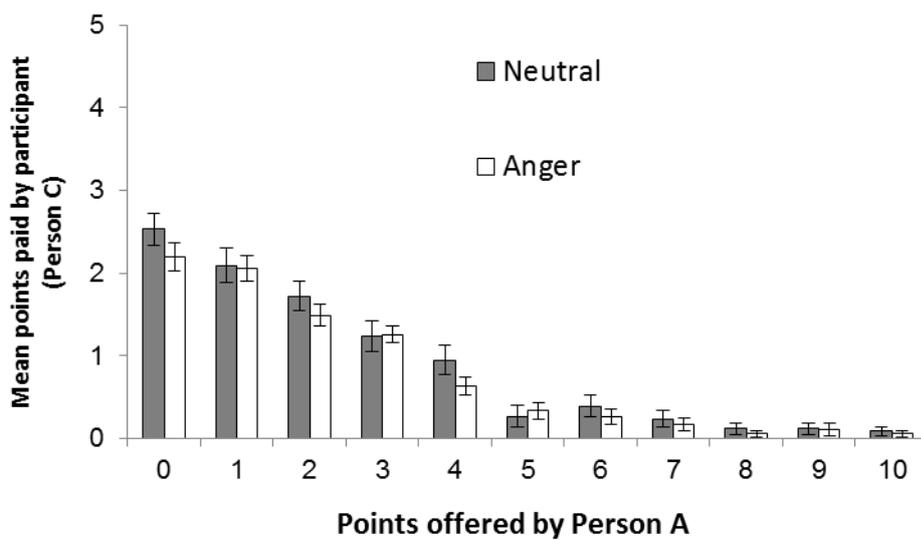
(a)

(b)

Figure 3. Study 2: Mean compensation rates (points paid to Person B) by Person C in an angry and neutral mood in (a) the wait condition and (b) the distraction condition. Error bars display standard errors.

(a)

(b)

Figure 4. Study 3: Mean compensation rates (points paid to Person B) by Person C in an angry and neutral mood in (a) the wait condition and (b) the distraction condition. Error bars display standard errors