



Gratitude increases third-party punishment

Jonathan Vayness, Fred Duong & David DeSteno

To cite this article: Jonathan Vayness, Fred Duong & David DeSteno (2019): Gratitude increases third-party punishment, *Cognition and Emotion*, DOI: [10.1080/02699931.2019.1700100](https://doi.org/10.1080/02699931.2019.1700100)

To link to this article: <https://doi.org/10.1080/02699931.2019.1700100>



Published online: 08 Dec 2019.



Submit your article to this journal [↗](#)



Article views: 128



View related articles [↗](#)



View Crossmark data [↗](#)

BRIEF ARTICLE



Gratitude increases third-party punishment

Jonathan Vayness, Fred Duong and David DeSteno

Department of Psychology, Northeastern University, Boston, MA, USA

ABSTRACT

Third-party punishment occurs when a perpetrator of a transgression is punished by another person who was not directly affected by the transgression (i.e. a third-party). Given gratitude's demonstrated ability to enhance both cooperation and the value people place on future-rewards, its capacity to increase third-party punishment – a phenomenon theorised to increase future cooperative behaviour – was investigated. In two experiments, participants were randomly assigned to experience one of three emotional states (i.e. gratitude, happiness, or neutrality) prior to making decisions about how much of a previous financial endowment they would spend to punish a person who transgressed against another at differing degrees within the context of a dictator game. As expected, punishment expenditures decreased for all participants as a dictator's decision became fairer. Of primary interest, however, participants who felt grateful, as compared to those who felt neutral or happy, engaged in significantly more third-party punishment across dictator splits that were not altruistic in nature.

ARTICLE HISTORY

Received 22 August 2019
Revised 25 November 2019
Accepted 26 November 2019

KEYWORDS

Gratitude; third-party punishment; emotion; morality

Third-party punishment (TPP) occurs when the perpetrator of a transgression is punished by an individual who is not the victim of the transgression and, therefore, not directly affected by it. As Fehr and Fischbacher (2004) demonstrated, not only will people often punish someone who treated another person unfairly, but they will even do so at cost to themselves. Although varying in degree, evidence for TPP has been shown to exist cross-culturally (Henrich et al., 2006).

People are believed to accept the immediate costs TPP requires due to its role in stabilising cooperative behaviour in social groups (Bendor & Swistak, 2001; Boyd, Gintis, Bowles, & Richerson, 2003; Fehr & Fischbacher, 2004; Henrich & Boyd, 2001). Historically speaking, it was likely that members of relatively small social groups would regularly engage in repeated exchange. Therefore, accepting a small cost to punish someone who transgressed against a third party could serve as a mechanism for helping to ensure that this person would not transgress again when his or her victim might be the present punisher. Indeed, people have been shown to infer an individual's mistreatment of others as an indicator for how

he or she would mistreat them and, consequently, have been shown to modulate their TPP actions accordingly (Delton & Krasnow, 2017; Krasnow, Delton, Cosmides, & Tooby, 2016).

At a proximate causal level, recent evidence confirms Fehr and Fischbacher's (2004) initial suspicion that moral emotional responses drive TPP. Moral outrage, or anger at unfair behaviour, has been shown not only to mediate TPP directly (Nelissen & Zeelenberg, 2009), but also to intensify it even when the anger is incidental to the current transgression (Gummerum, Van Dillen, Van Dijk, & López-Pérez, 2016). Interestingly, hedonically pleasant states like happiness, while producing less TPP when compared to negative states like anger, do not appear to reduce TPP when compared to affectively neutral states (Lane, 2017).

Such findings might suggest that emotions alter TPP as a function of valence, with only certain negative emotions being capable of enhancing punishment. Yet there are strong reasons to believe that the functional specificity, as opposed to valence, of states might better predict decisions to punish. By functional specificity, we mean the adaptive goals toward which

each emotion guides behaviour. Our past work demonstrates that gratitude, as opposed to general positive states like happiness, reduces temporal discounting, thereby increasing the value people place on future rewards relative to present ones (DeSteno, Li, Dickens, & Lerner, 2014; Dickens & DeSteno, 2016). As such, it might well lead people to be willing to pay a small cost in the present in order to help avoid a bigger loss in the future due to potential interactions with a person who continues to behave dishonestly.

This view suggests that those feeling grateful should be more motivated to attempt to ensure fair behaviour on the part of others via TPP. Indeed, grateful individuals evidence an enhanced neural valuation of altruistic acts (Karns, Moore, & Mayr, 2017) and, when given the opportunity, often engage in exchange behaviours where they willingly give greater, rather than equal, amounts to others (Bartlett & DeSteno, 2006; DeSteno, Bartlett, Baumann, Williams, & Dickens, 2010). Accordingly, they could be expected to hold a strong bias against inequitable, selfish acts on the part of anyone.

It is important to note that this prediction reflects an incidental effect of gratitude as opposed to one wherein gratitude derives from direct interaction with the two targets comprising situations of TPP (i.e. victim and transgressor). Thus, increases in gratitude can be expected to correspondingly increase TPP even though nothing about the moral transgression itself evoked gratitude. Rather, current feelings of gratitude on the part of an observer can be expected to influence the interpretation and valuation of the events taking place within the context of a separate moral transgression. To examine this hypothesis, we conducted two experiments – the second a replication of the first – being careful to contrast the effects of gratitude not only to that of a neutral state, but also to a more general positive one, happiness, to rule out the possibility that any resulting effect could be explained solely by valence. In accord with Lane (2017), no differences in punishment should be expected to emerge between those feeling neutrally and those feeling happy.

Study 1

Method

Open practices statement

Materials and data for this experiment can be accessed at <https://osf.io/2cuhp/>. This study was not preregistered.

Ethical approval

All procedures performed in the studies reported here were reviewed by and conducted in accordance with the ethical standards of Northeastern University Institutional Review Board (reference # 17-12-03), and are in accord with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Participants

The calculated sample size was based on the mean effect size from two previously published experiments that compared the influence of gratitude to that of happiness and neutral states on costly prosocial behaviour (Bartlett & DeSteno, 2006; DeSteno et al., 2014): Cohen's $f = .348$. Using G*Power, we determined that 84 participants were required to achieve a power = .80. Given this requirement, we set 84 as the lower bound for a sufficient sample and attempted to recruit as many participants as possible given the time and resources available for conducting this experiment. The final sample consisted of 96 individuals (63 women, 33 men, $M_{age} = 19$, $SD_{age} = 1.02$ years) drawn from the university participant pool who we randomly assigned to one of the three emotion-induction conditions (prospective power = .86).

Procedure

Upon arriving at the lab, participants were seated in individual cubicles containing a desktop computer. They were informed that they would complete two separate tasks: one involving memory and one involving an economic game. In actuality the memory task functioned as an emotion induction and the economic game as a measure of third-party punishment.

Emotion induction. Following the procedures used by DeSteno et al. (2014), participants were informed that they would be randomly assigned to recall and write about a certain type of event from their past. Depending on condition, participants' computers instructed them to write about a time they were grateful, happy, or about the events of a normal day (i.e. neutral condition). After the allotted 5 min has passed, participants next completed an emotion manipulation check wherein, using a 5-point scale, they indicated the degree to which each of 17 different feeling items described their current emotional state. Gratitude was calculated as the mean of: *grateful*, *appreciative*, and *thankful*

(Cronbach's $\alpha = .94$). Happiness was calculated as the mean of *happy*, *content*, and *pleasant* (Cronbach's $\alpha = .85$).

Third-party punishment. After completing the emotion induction, participants began the third-party punishment game, which took the form of a third-party dictator game (TP-DG). The TP-DG followed a slightly modified version of that used in Fehr and Fischbacher (2004), which employed a basic two-player dictator game with an additional third player who had the option of punishing the dictator for his or her decisions. All participants were told that they had been assigned to the third player role.

The game was played with points at a conversion rate of 10 points being equal to \$1.00. Participants were told that in a previous experimental session that day, one player (the dictator) had been given 100 points and had the option to split those points using increments of 10 with another player who was given no initial endowment (i.e. player 2). Participants were then endowed with 50 points (\$5) which they could keep, or upon seeing how dictators split their endowment with player 2, could spend to deduct points from the dictator. Every point participants used from their own endowment for this purpose would cause the dictator to lose three points (i.e. 30¢) from his or her remaining point total while reducing their own endowment by 10¢.

After reading these instructions, participants advanced to a screen in which they were told that before seeing the actual split the dictator chose, they would have to indicate how many deduction points they would assign given all of the possible ways that the dictator could have split the endowment with a partner (i.e. dictator kept 100 and gave 0, dictator kept 90 and gave 10 ... dictator kept 0 and gave 100).

Participants were then presented with a single screen on which the 11 different split options were presented, told to consider the options, and then to indicate their punishment decisions, if any, for each. For the purpose of studying third-party punishment, we intended to examine responses to the range of splits characterising complete selfishness (i.e. 100/0) to equitable sharing (i.e. 50/50), as these were the cases in which elevated punishment for gratitude could be expected given that grateful participants show greater positivity toward and engagement in non-selfish behaviours. Nonetheless, we also

examined participants' responses to altruistic splits to as a check to ensure that no substantial punishment was occurring therein.

After indicating their decisions, participants were shown what they believed to be the dictator's decision, though in actuality this decision was selected randomly by the computer. The amount of points they had indicated to be used from their own endowment were then removed and the remainder of points were then converted to money that participants kept.

Results

Manipulation check

We submitted participants' responses to the gratitude and happiness scales to a 3 (Emotion Condition: Gratitude, Happiness, Neutral) X 2 (Emotion Intensity: Gratitude vs. Happiness) mixed ANOVA. As expected, the manipulations produced distinct emotional states as indicated by a significant interaction, $F(2, 93) = 9.79$, $p < .001$. A one way ANOVA on gratitude intensity proved significant, $F(2, 93) = 20.44$, $p < .001$, with Fisher's LSD comparisons confirming that those who completed the gratitude manipulation ($M = 4.58$, $SD = 0.55$) reported significantly greater intensities of gratitude than did those who completed the neutral ($M = 3.02$, $SD = 1.29$) and happiness ($M = 3.28$, $SD = 1.17$) manipulations.

In a similar vein, a one way ANOVA on happiness intensity emerged, $F(2, 93) = 6.79$, $p = .002$, showing that those who completed the happiness ($M = 3.75$, $SD = 0.90$) and gratitude ($M = 3.92$, $SD = 0.79$) manipulations reported greater intensities of happiness than did those who completed the neutral ($M = 3.17$, $SD = 0.87$) manipulation. The fact that those in the gratitude and happiness conditions reported similar levels of happiness but divergent levels of gratitude is to be expected, as the experience of gratitude itself usually co-occurs with positive feelings toward a benefactor for her actions (e.g. giving one a gift, helping one out of a jam). Elevation of feelings of gratefulness in the presence of similar levels of happiness is what defines a distinct state of gratitude that has been shown to lead to different behavioural outcomes from that of happiness alone (Bartlett & DeSteno, 2006; DeSteno et al., 2014; DeSteno, Duong, Lim, & Kates, 2019).

Third party punishment

To assess how the different emotions and nature of the dictator's decisions shaped third-party

punishment, we conducted a 3 (Emotion Condition: Gratitude, Happiness, Neutral) X 6 (Split: 100/0 to 50/50) mixed ANOVA. Because both Mauchly's test of sphericity for the repeated factor and Levene's median-based statistic for the between-groups factor were significant (p 's $\leq .001$), it became clear that heterogeneity of variance and covariance in punishment amounts existed across conditions. Consequently, we report Greenhouse-Geisser or similar corrected statistics for inferential tests where relevant.

As expected, we found no evidence that the different emotions altered the rate at which punishment changed as splits became fairer, $F_{\text{interaction}}(2.99, 138.85) = 1.31$, $p = .34$, partial $\eta_p^2 = 0.03$, 95% CI .000–.088. Rather, the significant main effect for the dictator's splits, $F(1.49, 138.85) = 105.88$, $p < .001$, $\eta_p^2 = 0.53$, 95% CI .415–.615, showed that TPP followed a similar linear decline for all emotion groups as decisions became less selfish, $F_{\text{linear trend}}(1, 93) = 131.59$, $p < .001$, $\eta_p^2 = .59$, 95% CI .477–.661 (See Figure 1).

Our primary prediction centred on whether people feeling gratitude would engage in greater third-party punishment than would others. A contrast on the emotion factor that allowed for heterogeneity of variance and compared punishment expenditures of grateful participants to those of happy and neutral participants confirmed this prediction, Welch's $t(42.02) = 2.32$, $p = .025$, $\eta_p^2 = 0.08$, 95% CI .012–.171. The residual for the contrast was not

significant, $F < 1$, thereby indicating that the identified pattern of mean differences accounted for all systematic variation among the three group means. Whereas those feeling grateful paid \$1.26 of their initial \$5 (i.e. 12.60 points on average) on average across split conditions to punish the dictator, those feeling happy and neutral paid lesser amounts that did not significantly differ from each other (69¢ and 88¢, respectively; 6.95 and 8.76 points on average). Put another way, participants feeling grateful were willing to spend approximately 10% more of their profits to punish transgressors than were their neutral or happy peers.

Although not directly relevant to the primary questions and analyses presented here, we also examined whether any differences in punishment behaviour for splits where the dictator acted altruistically (i.e. 40/60 to 0/100) existed. As expected, we found no systematic effects either as a result of the split ratio ($p = .073$) or emotion of the potential punisher ($p = .192$). The mean points paid on average across these conditions was 2.22 (95% CI 1.39–3.05), with nonzero values likely indicating a small propensity for antisocial punishment (see Herrmann, Thoni, & Gächter, 2008). As one might expect, a mixed ANOVA on the full set of splits between the dictator and player (100/0 to 0/100) as a function of emotion produced a significant Emotion Condition X Split interaction, $F(4.54, 210.98) = 2.47$, $p = .039$, $\eta_p^2 = .05$, 95% CI .013–.054.

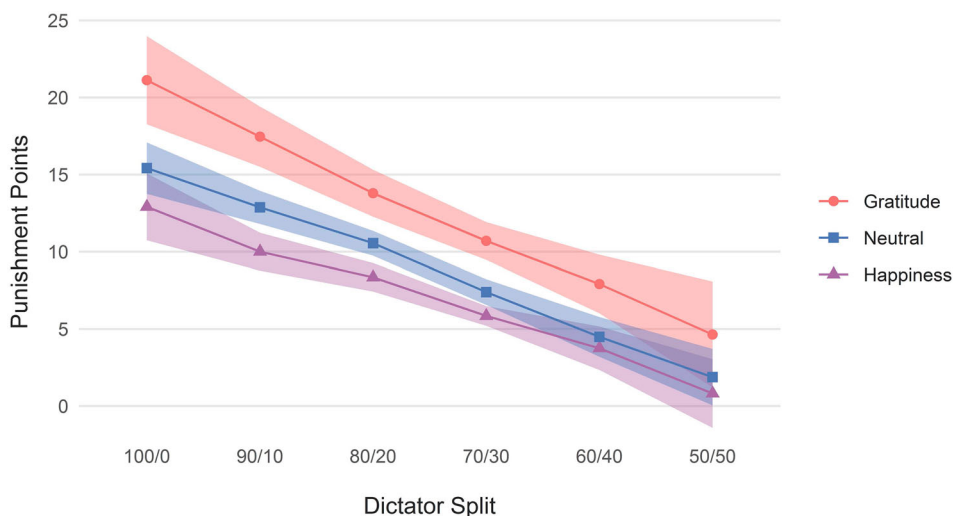


Figure 1. Points spent to punish the transgressor as a function of emotional state and the dictator's split is Study 1. Shaded areas represent 95% confidence intervals.

Study 2

Method

Open practices statement

Materials and data for this experiment can be accessed at <https://osf.io/2cuhp/>. Although this study was not preregistered, findings from Study 1 were posted at <https://psyarxiv.com/29ebh/> prior to the conduct of Study 2 (see Version 1 of preprint). Study 2, with the exception of the minor modifications to participant endowments and punishment costs noted below, followed the same methodological and analytic procedures as those used in Study 1. As such, Study 2 is an almost exact replication of Study 1, with all analytic decisions constrained to be the same.

Participants

The size of the primary effect of interest from Study 1 was used to calculate the required sample size for Study 2: Cohen's $f = .294$. Using G*Power, we determined that 120 participants were required to achieve a power = .80. However, as we would be recruiting from Amazon MTurk, we anticipated the loss of some data due to quality reasons. Therefore, we set recruitment at 150 participants. Ultimately seven participants were removed for a failure to adhere to instructions regarding the emotion manipulation: they produced responses for the autobiographical recall task that were not focused on the assigned topic. The final sample consisted of 143 individuals (59 women, 83 men, 1 not reported, $M_{age} = 35$, $SD_{age} = 10.67$ years) drawn from the pool of Amazon MTurk workers (prospective power = .87).

Procedure

The measures and procedure for Study 2 were nearly identical to those used in Study 1. The minor differences involved the amount of money with which participants were endowed, the financial point system available for punishment, the possible dictator splits that were presented, and the fact that participants were also paid \$1.00 as compensation for completing the study on MTurk irrespective of any TPP choices made. The third-party punishment game followed the same format as before, with the exception that participants were endowed with 50 points, with 1 point now equalling 1¢, for an endowment of 50¢, and the fictional dictator with 100 points. As a result, the maximum bonus a participant could receive was

50¢. Note that the ratios of the initial endowments with respect to costs to punish are consistent across studies. Finally, because we were primarily interested in selfish vs. equal splits (as opposed to altruistic ones) and because, as expected, no effects of gratitude emerged on altruistic splits in Study 1, here we only presented participants with splits ranging from 100/0 to 50/50 percentages.

Results

Manipulation check

As in Study 1, we submitted participants' responses to the gratitude and happiness scales to a 3 (Emotion Condition: Gratitude, Happiness, Neutral) X 2 (Emotion Intensity: Gratitude vs. Happiness) mixed ANOVA. The manipulations again produced distinct emotional states as indicated by a significant interaction, $F(2, 140) = 9.41$, $p < .001$. A one way ANOVA on gratitude intensity proved significant, $F(2, 140) = 10.39$, $p < .001$, with Fisher's LSD comparisons confirming that those who completed the gratitude manipulation reported significantly greater intensities of gratitude ($M = 4.10$, $SD = 0.86$) than did those who completed the neutral ($M = 3.04$, $SD = 1.17$) and happiness ($M = 3.36$, $SD = 1.45$) manipulations.

In a similar vein, a one way ANOVA on happiness intensity emerged, $F(2, 140) = 8.60$, $p < .001$, showing that those who completed the happiness ($M = 3.58$, $SD = 1.19$) and gratitude ($M = 3.31$, $SD = 0.93$) manipulations reported greater intensities of happiness than did those who completed the neutral ($M = 2.73$, $SD = 1.00$) manipulation. The fact that those in the gratitude and happiness conditions reported similar levels of happiness but divergent levels of gratitude was again to be expected as it was in Study 1. Heightened feelings of gratefulness in the presence of similar levels of happiness defines a distinct state of gratitude (Bartlett & DeSteno, 2006; DeSteno et al., 2014; DeSteno et al., 2019).

Third party punishment

As in Study 1, we conducted a 3 (Emotion Condition: Gratitude, Happiness, Neutral) X 6 (Split: 100/0 to 50/50) mixed ANOVA on third party punishment decisions. The assumption of sphericity was again violated, so we present Greenhouse-Geisser adjustments for all tests involving the repeated factor. As expected, we found no evidence that the different emotions altered the rate at which punishment changed as splits became fairer, $F_{interaction}(3.83, 268.25) = 1.10$, p

$= .36$, $\eta_p^2 = 0.02$, 95% CI .000–.055. Replicating Study 1, a linear trend contrast again showed that TPP followed a linear decline for all emotion groups as decisions became less selfish, $F(1, 140) = 23.30$, $p < .001$, $\eta_p^2 = .14$, 95% CI .064–.231 (See Figure 2).

Our primary prediction centred on whether people feeling gratitude would engage in greater third-party punishment than would others. A contrast on the emotion factor that compared punishment expenditures of grateful participants to those of happy and neutral participants again confirmed this prediction, $t(140) = 2.09$, $p = .038$, $\eta_p^2 = 0.03$, 95% CI .001–.090.¹ The residual for the contrast was not significant, $F < 1$, thereby indicating that the identified pattern of mean differences accounted for all systematic variation among the three group means. Whereas those feeling grateful paid 17¢ (17.36 points on average) of their initial 50¢ on average across split conditions to punish the dictator, those feeling happy and neutral paid lesser amounts that did not significantly differ from each other (13¢ and 12¢, respectively; 12.56 and 11.65 points on average). Similar to Study 1, participants feeling grateful were willing to spend approximately 9% more of their profits to punish transgressors than were their neutral or happy peers.

General discussion

The results reported here clearly show that gratitude, as differentiated from the more general positive

state of happiness, increases TPP. Although this fact might seem surprising at first, it fits well with related findings showing that moral emotions can engender behaviours that, though seemingly problematic on their own, actually serve a prosocial function (Lupoli, Jampol, & Oveis, 2017; Ng et al., 2017). We feel it important to stress that we are in no way implying that gratitude is the only morally-toned emotion which might heighten TPP. As noted, negative feelings of moral outrage increase TPP (Gummerum et al., 2016; Jordan & Rand, *in press*; Nelissen & Zeelenberg, 2009). Other positive states can as well. For example, recent evidence suggests that compassion felt toward victims increases TPP (Pfattheicher, Sassenrath, & Keller, 2019). These moral emotions nudge people toward acts designed to reinforce cooperation through strengthening relationship bonds, norms for moral behaviour, or signals that a person can be counted on to favour fair outcomes.

In the present case, we believe that gratitude likely pushed people to punish others who did not act fairly in an effort to enhance their future trustworthiness. Such behavioural shaping is central to third-party observers' outcomes as, historically speaking, they were likely to interact with transgressors at a later time. However, it should be noted that other potential explanations for the effect of gratitude are equally viable.

One such explanation centres on gratitude's propensity to build social bonds by nudging people to act in less selfish ways. Specifically, gratitude has

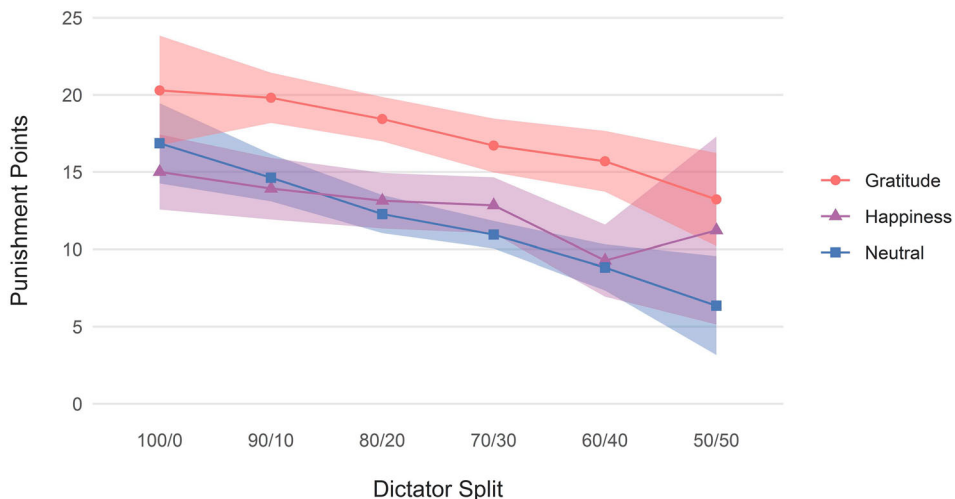


Figure 2. Points spent to punish the transgressor as a function of emotional state and the dictator's split in Study 2. Shaded areas represent 95% confidence intervals.

been found to foster cooperation and related prosocial behaviours (Algoe, 2012; Bartlett & DeSteno, 2006; DeSteno et al., 2010; DeSteno et al., 2019; Ma, Tunney, & Ferguson, 2017). As such, it might well lead people to wish to punish bad actors according to a just deserts perspective wherein social wrongs must be righted (Carlsmith, Darley, & Robinson, 2002). Put another way, because gratitude enhances prosocial motivations, it might also increase the degree to which people punish antisocial behaviour, either because the severity of the current transgression or desire to prevent subsequent transgressions is higher.

Another explanation centres on social norms. Because gratitude motives people to maintain norms in an effort to stabilise group harmony (Ng et al., 2017), it could also be expected to make people more sensitive to norm violations. That is, it might well push people to weight present or future violations of moral norms more heavily, thereby increasing their willingness to engage punishment to remedy or reduce conflict.

In conclusion, we wish to make a final observation about the conditions under which gratitude's effect on TPP would be most likely to occur. While we agree that witnessing a transgression would be unlikely to evoke gratitude, re-experiencing feelings of gratitude when thinking about or seeing previous benefactors could be expected to enhance TPP upon learning of their victimisation. In this way, gratitude, over time, could continue to protect those who have behaved honestly in the past.

Note

1. Although the use of contrast analyses was planned *a priori*, and reflects our usual analytical approach in studies on gratitude that utilise both a neutral and happiness control (e.g. Bartlett & DeSteno, 2006; DeSteno et al., 2014), additional evidence that gratitude elevates punishment compared to a neutral affective state can come from direct comparisons between these two conditions. In Study 1, a t-test shows the difference between the gratitude and neutral conditions to approach the standard level of significance ($p = .058$). In Study 2, which uses a larger sample size with a power analysis based on the effect size for the contrast obtained in Study 1, the difference between the gratitude and neutral conditions is significant ($p = .035$). Combining these two findings in a mini-meta analysis using Stouffer's procedure confirms a reliable difference in punishment for grateful vs. neutral participants ($p = .016$). Note that the power analyses used to determine the sample size in both studies utilised an effect size based on contrast

analyses as opposed to an omnibus test, as this was planned analytic tack.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Algoe, S. B. (2012). Find, remind, and bind: The functions of gratitude in everyday relationships. *Social and Personality Psychology Compass*, 6, 455–469.
- Bartlett, M. Y., & DeSteno, D. (2006). Gratitude and prosocial behavior: Helping when it costs you. *Psychological Science*, 17(4), 319–325.
- Bendor, J., & Swistak, P. (2001). The evolution of norms. *American Journal of Sociology*, 106(6), 1493–1545.
- Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6), 3531–3535.
- Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology*, 83, 284–299.
- Delton, A. W., & Krasnow, M. M. (2017). The psychology of deterrence explains why group membership matters for third-party punishment. *Evolution and Human Behavior*, 38(6), 734–743.
- DeSteno, D., Bartlett, M. Y., Baumann, J., Williams, L. A., & Dickens, L. (2010). Gratitude as moral sentiment: Emotion-guided cooperation in economic exchange. *Emotion*, 10(2), 289–293.
- DeSteno, D., Duong, F., Lim, D., & Kates, S. (2019). The grateful don't cheat: Gratitude as a fount of virtue. *Psychological Science*, 30, 979–988.
- DeSteno, D., Li, Y., Dickens, L., & Lerner, J. S. (2014). Gratitude: A tool for reducing economic impatience. *Psychological Science*, 25(6), 1262–1267.
- Dickens, L., & DeSteno, D. (2016). The grateful are patient: Heightened daily gratitude is associated with attenuated temporal discounting. *Emotion*, 16(4), 421–425.
- Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2), 63–87.
- Gummerum, M., Van Dillen, L. F., Van Dijk, E., & López-Pérez, B. (2016). Costly third party interventions: The role of incidental anger and attention focus in punishment of the perpetrator and compensation of the victim. *Journal of Experimental Social Psychology*, 65, 94–104.
- Henrich, J., & Boyd, R. (2001). Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, 208(1), 79–89.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., ... Ziker, J. (2006). Costly punishment across human societies. *Science*, 312(5781), 1767–1770.
- Herrmann, B., Thoni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, 319, 1362–1367.
- Jordan, J. J., & Rand, D. G. (in press). Signaling when no one is watching: A reputation heuristics account of outrage and punishment in one-shot anonymous interactions. *Journal of Personality and Social Psychology*.

- Karns, C. M., Moore, W. E., & Mayr, U. (2017). The cultivation of pure altruism via gratitude: A functional MRI study of change with gratitude practice. *Frontiers in Human Neuroscience*. doi:10.3389/fnhum.2017.00599
- Krasnow, M. M., Delton, A. W., Cosmides, L., & Tooby, J. (2016). Looking under the hood of third-party punishment reveals design for personal benefit. *Psychological Science*, 27(3), 405–418.
- Lane, T. (2017). How does happiness relate to economic behaviour? A review of the literature. *Journal of Behavioral and Experimental Economics*, 68, 62–78.
- Lupoli, M. J., Jampol, L., & Oveis, C. (2017). Lying because we care: Compassion increases prosocial lying. *Journal of Experimental Psychology: General*, 146, 1026–1042.
- Ma, L. K., Tunney, R. J., & Ferguson, E. (2017). Does gratitude enhance prosociality? A meta-analytic review. *Psychological Bulletin*, 143, 601–635.
- Nelissen, R. M., & Zeelenberg, M. (2009). Moral emotions as determinants of third-party punishment: Anger, guilt and the functions of altruistic sanctions. *Judgment and Decision Making*, 4(7), 543.
- Ng, J. W. X., Tong, E. M. W., Sim, D. L. Y., Teo, S. W. Y., Loy, X., & Giesbrecht, T. (2017). Gratitude facilitates private conformity: A test of the social alignment hypothesis. *Emotion*, 17, 379–387.
- Pfafftheicher, S., Sassenrath, C., & Keller, J. (2019). Compassion magnifies third-party punishment. *Journal of Personality and Social Psychology*, 117(1), 124–141.