Is peer punishment needed under threat?
Christopoulou, Anna

Universiteit Leiden

Faculteit der Sociale Wetenschappen

Is peer punishment needed under threat?

Anna Christopoulou

In collaboration with Laura Cirigliano and Coert Pen

Master thesis Psychology, specialization Social and Organizational Psychology
Institute of Psychology
Faculty of Social and Behavioural Sciences – Leiden University
Date: 18/02/2022
Student number: s3001431
First examiner of the university: Dr. Maria Lojowska
Second examiner of the university: Dr. Jörg Gross

Abstract

People are exposed to a wide range of aversive conditions, including climate change and pandemics, which can have a profound impact on individual wellbeing and communal functioning. While there is evidence suggesting that exposure to threat and the prospect of punishment promote cooperation, the necessity of punishment in cooperation's maintenance under threat is largely unexplored. In the current study, we examined whether the presence of threat requires less punishment to maintain cooperation. In a laboratory experiment individuals in groups of three (N=60) were exposed (or not) to the threat of electric shocks while deciding how much to contribute to the common pool. Additionally, half of the tested groups (N=30) were subjected to a peer-punishment procedure in which individuals had the option to deduct money (MU's) from other group members' accounts. Heart rate and skin conductance were continuously measured while participants were exposed to the aforementioned procedure. In comparison to the no-threat condition, the threat of shock resulted in a neurophysiological freezing response characterized by a reduction in heart rate and an increase in skin conductance. We find that in contrast to our expectations, threat by itself does not promote cooperation and punishment is needed in cooperation's maintenance. Overall, our results suggest that the presence of threat does not diminish the role of punishment in increasing cooperation. Our results have implications for policy interventions designed to sustain cooperation under threat.

Keywords: threat, punishment, cooperation, public goods game

Introduction

Aversive conditions such as pandemics, climate change, and economic shortages pose threats that, while having negative consequences on both individual and societal levels, can be mitigated by cooperation (Boyd & Richerson, 2009; Fehr & Fischbacher, 2004; Rand & Nowak, 2013). But what makes people remain cooperative in threatening situations? There are findings indicating that the experience of threat can promote cooperation (Dezecache et al., 2017; Shentu et al., 2018, Lojowska et al., 2022). Another means shown to facilitate cooperation is peer-punishment (Lohse & Waichman, 2020). Peer-punishment has been found to increase cooperation and reduce free-riding. Specifically, people are willing to pay a personal cost in order to penalise those who benefit from the public good without contributing themselves (Fowler, 2005; Hauert et al., 2007). Compared to threat-induced cooperation, peer punishment is, however, costly because it involves a waste of resources (e.g. waste of time or money) not only for the punished but also for the punisher. The aim of the current study was to investigate whether, in the presence of threat, cooperation would be maintained with less punishment, provided a facilitating effect of threat on cooperation.

Throughout the evolution, humans have collectively fought threats such as contagious diseases and natural disasters by cooperative efforts (Hauser et al., 2014; Milinski et al., 2006, 2008). Although such threats have profound negative personal and social consequences, people can achieve more synergistically than individually (Kallhoff, 2014; Mawson, 2005; Mobbs et al., 2020). In pandemics, for instance, investing in public goods, such as in public healthcare is more beneficial than investing in individual solutions such as buying a large number of very expensive antibiotics. However, human cooperation is rather perplexing. That is because public goods necessitate everyone's eagerness to contribute to their group, whereas consumption is available to everyone, not limited to contributors (Dietz et al., 2003; Hardin, 2009). Thus, this aspect of non-excludability increases the chance that people will try to benefit individually by enjoying the benefits while not contributing to the common good. As a result, the public good is often underprovided. The question is how cooperation can be sustained when social and individual interests clash. To explain human cooperation, several theories have been proposed. For instance, the theory of kin selection proposes that cooperation is preserved among genetically related individuals (Hamilton, 1964). Theories of direct reciprocity suggest that selfish motives enhance cooperation in long-term interactions (Dal Bo, 2005; Trivers, 1971). The theories of indirect reciprocity imply that cooperation emerges in larger groups because cooperators can establish reputation (Leimar & Hammerstein, 2001; Nowak & Sigmund, 1998). Yet, cooperation can break down among unrelated people, in single interactions, and when reputational gains are small or absent. Since human cooperation is fundamentally fragile, communities seek ways to sustain it.

Humans typically respond to threatening situations by undergoing neurophysiological changes that influence decision-making (Fanselow, 1994; Öhman et al., 2001). At the physiological level, defensive responses to threat manifest as sympathetically controlled increases in skin conductance and parasympathetically driven decreases in heart rate (Gladwin et al., 2016; Mackersie & Calderon-Moultrie, 2016; Roelofs, 2017). This physiological state is typical of the physiological "freezing" response. Freezing triggers a state of hypervigilance and enhanced sensory processing, which allows the evaluation of threat-relevant cues, resulting in optimal threat responses (Blanchard et al., 2011; Lang et al., 2013; Lojowska et al., 2018; Sokolov & Cacioppo, 1997). Freezing, as a reaction to threat, helps the person to adapt and improve its chances of survival. As a result of their heightened emphasis on self-protection in the freezing state, humans may prioritize personal interests over the interests of others (Engelmann et al., 2019; FeldmanHall et al., 2015). For example, Engelmann and colleagues (2019), found that threat exposure influences the neural circuity related to emotion processing and as a result, at a behavioural level reduces trust among people.

Whereas previous research has contributed to explaining how people react to threat exposure individually, the majority of threats humans encounter are, in fact, collective in nature, and joint actions are frequently required to avoid them (e.g. climate change, pandemics). The key element of collective threats is that individual responses and decisions have an impact on not only themselves but also on others (Gross & Dreu, 2019; Van Lange & Rand, 2022). Indeed, it has been found that cooperation among people breaks down over time due to the conflict between self-interest and collective interest (Fehr & Gächter, 2001; Lehtonen & Jaatinen, 2016; van Dijk & De Dreu, 2021). However, when individual survival is at stake, more collective responses are observed among group members exposed to aversive conditions (Hamilton, 1971; Ioannou, 2021; Lehtonen & Jaatinen, 2016). Specifically, aggregation in response to predation is commonly observed in various animals such as fish and birds (Beauchamp, 2004; Hoare et. al., 2004), which according to the “safety in numbers” principle helps to increase the survival of an individual (Ioannou, 2021; Lehtonen & Jaatinen, 2016). Collective responses and prosocial decision making in threatening situations have also been observed in humans. People may “stick together” and take collective action to protect themselves (Calo-Blanco et al., 2017; Lojowska et al., 2022; Mawson, 2005; Tedeschi et al., 2020; Vieira et al., 2020). For example, Tedeschi and colleagues (2020), found that when there is a chance of threat, people tend to band together and choose groups with a larger number of members. In a real-life example, Vieira and her colleagues (2020) examined how the threat caused by the COVID-19 pandemic influenced people’s helping behaviour. The results revealed that the experience of anxiety triggered by the experience of ongoing pandemics (i.e., COVID-19) was associated with more altruistic behaviour, such as helping strangers or donating blood. Finally, in a laboratory experiment, Lojowska and colleagues (2022) tested the effect of anticipatory threat state on cooperation in small groups. They used a public goods game where participants decided in multiple trials their (monetary) contributions to the common pool, either under the threat of electric shocks or in a safe condition. The results revealed that whereas in non-threatening conditions, a typical pattern of time-dependent reduction in cooperation was observed, a sustained level of high cooperation was observed during the threat condition Overall, these data support a view that threat promotes affiliation, and that humans tend to act more prosocially during collective exposure to threats.

Peer punishment has been recommended as a strategy for maintaining cooperation, mainly by reducing free riding. Peer punishment refers to people’s tendency to punish others, primarily, for their lack of cooperation. Cooperation gradually deteriorates due to the risk of cooperators being exploited by those who profit from public assets without contributing (free-riders). However, if those who profit from others' cooperation are penalized, free-riding becomes less attractive and, thus, cooperation may persist (Balafoutas et al., 2014; Fehr & Schurtenberger, 2018; Lohse & Waichman, 2020; Molleman et al., 2019) Altruistic punishment, although beneficial for cooperation, might be viewed as a "double-edged sword": it helps to maintain cooperation, but is costly for the punisher. For instance, some people could argue with someone who parks their car in a disabled parking space. On a personal level, there is no benefit from this activity, but it can be costly due to time-wasting. The motivational factors of altruistic punishment have been widely studied (Li et al., 2021; Nikiforakis, 2010). Personality characteristics, such as inequality aversion, which refers to people's inclination to take actions that promote fairness and eliminate inequality, prosociality, which refers to people's tendency to care for the well-being of others, and altruism, which refers to people's proclivity to treat others with unreserved generosity, explain why some people are eager to punish others (Engel, 2014; Fehr & Fischbacher, 2003; Fehr & Gächter, 2002; Fehr & Schmidt, 2010; Wang et al., 2021). Overall, despite its puzzling nature, in a number of experimental research punishment has been shown to promote cooperation and diminish free-riding, owing to people's eagerness to punish those who take advantage of others' cooperation, even if it is costly on a personal level (Boyd & Richerson, 2009; de Quervain et al., 2004; Yamagishi, 1986).

Together, both threat and peer punishment have been shown to promote cooperativeness. The above evidence, however, raises a question of whether the presence of threat requires less punishment to maintain cooperation, given the waste of resources associated with punushment.

The aim of this study is to examine whether an anticipatory threat state can reduce the need for peer punishment in cooperation maintenance. To this end, we used a public good game with punishment. Participants played the game under the threat of electric shocks and safe condition (no shocks). Additionally, half of the groups were given a punishment option, i.e., to deduct Deduction Points (DPs) from other participants' accounts. We hypothesised that if the facilitating effects of threat reduce the need for peer- punishment, less Deduction Points will be deducted in threat compared to safe condition in punishment condition. Additionally, we expected to replicate previous findings on the enhancing effects of threat on cooperation (Lojowska et al., 2022), especially in the control condition. We also expected, and in line with previous studies, to find overall larger cooperation during punishment versus control (no-punishment) condition (Fehr & Gächter, 2002; Fehr & Fischbacher, 2003). Lastly, to examine whether our threat manipulation was successful in evoking a typical physiological response pattern under threat, we measured participants' heart rate and skin conductance. We expected that if anticipatory threat triggers sympathetic and parasympathetic activation, we should observe an increase in skin conductance and a decrease in heart rate, respectively.

Materials and Methods

Participants and ethics

Using G Power 3 software (Faul et al., 2007), we calculated that to achieve results at adequate power ($1 – \beta > 0.8$) and medium effect size (η^2=.06), a sample size of 176 participants was required to test. Therefore, a total of N = 60 three-person groups were collected (30 groups for each between-subject condition). Of the 180 tested participants, 131 were females and the average age of all participants was 21.6 years old. All groups were mixed gender. As assessed using the Social Value Orientation task (SVO, Murphy et al., 2011), 153 were "prosocial", 26 were "individualists" and 1 was "altruist". 0.77% of all trials were removed from the analyses because participants did not contribute actively in these trials, and in which contributions were decided by the computer. Trials involving shocks (i.e. 10% of all trials) were also removed from the analysis since we were primarily interested in anticipatory threat states rather than threat exposure itself.

The characteristics of the sample were the following: healthy participants aged between 18 and 35 years, no history of neurological, psychiatric, and cardiovascular conditions, sufficient understanding of written and spoken English and no use of psychotropic drugs within the past two weeks, no pregnancy (or doubt of being pregnant).

The study was approved by the ethics committee of the Faculty of Social Science of Leiden University, The Netherlands (ethics approval number: 2021-05-03-M. Lojowska-V2-3170). All participants provided written informed consent before the experiment and upon completion debriefed and paid and received financial compensation or course credits for their participation (i.e., 3.50 euro or 2-course credits). In addition, participants earned extra money with the decisions they made during the experiment (ranging from 7.07 to 17.26 euro). To preserve anonymity, payment was computed individually paid immediately after the end of the task by mobile bank transfer. No deception was involved in this experiment.

Shock administration and physiological measurements

To examine our hypothesis of the effect of threat on group cooperation, we induced an anticipatory threat state. The anticipatory threat state was operationalized through the chance of receiving unpleasant, but not painful electric shocks, during the public goods game. Electric shocks were delivered transcutaneously through the participant's fourth and fifth distal phalanges of a dominant hand using Digitimer Constant Current Stimulator DS5 or DS7 (www.digitimer.com) and standard Ag/AgCl electrodes. The duration of the electric stimulation was 200 ms, with a 50Hz repetition of 10 pulses. The intensity of electric shocks will vary between 1.2mA (level 1) to 10mA. The intensity of the shocks used in the experiment was adjusted at the individual level to ensure that the shocks were unpleasant, but not painful. Shock calibration was performed using a standard shock calibration procedure comprising approximately 5 shock presentations. After each shock administration, its intensity was adjusted according to the participant's verbal reports of its unpleasantness on a scale from 1 (not unpleasant) to 5 (very unpleasant). When the intensity was rated as 5, participants were asked a follow-up question about whether the shock was painful. If they responded with a no, the current shock level was subsequently used in the experiment. If they responded with a yes, the shock level was reduced, and only a shock level rated as non-painful was subsequently applied in the study. The shock intensity obtained during this shock calibration procedure was subsequently used used in the public goods game.

To objectively assess whether our threat manipulation was successful, we recorded heart rate (HR) and skin conductance (SC). HR was used for offline assessment of the parasympathetically controlled heart rate deceleration, i.e., physiological index of freezing, whereas SC was used as an index of sympathetic activity. HR and SC were acquired throughout the task using a BIOPAK MP 150 system (Biopak Systems, Goleta, CA, USA). The sample rate was set to 1,000 Hz. HR was measured using disposable ECG electrodes attached to the participants' chests. Skin conductance data were collected with two standard Ag/AgCl electrodes attached to the second and third distal phalanges of the participant's non-dominant hand **(Figure 1)**.

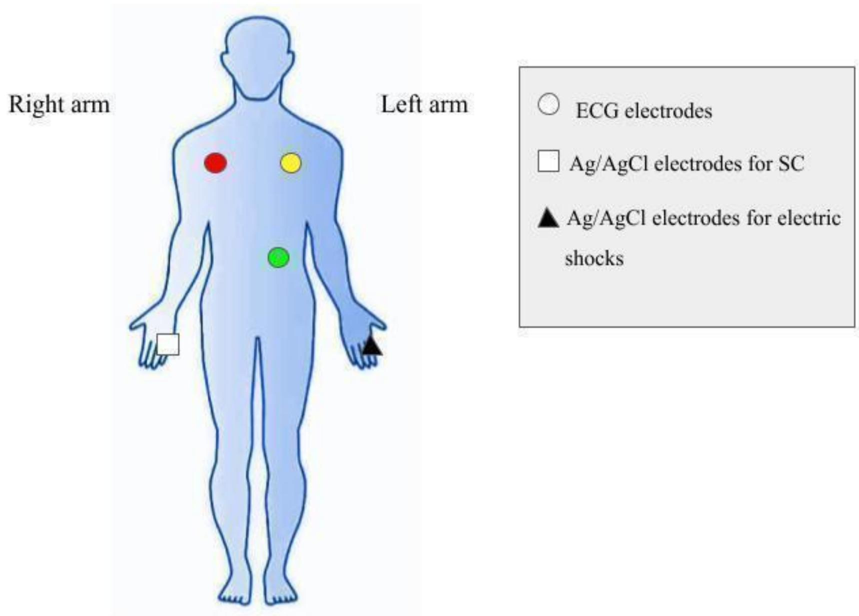

Figure 1. Physiological measurements. Heart rate (HR) was measured by disposable ECG electrodes attached to the participants' chests (sample rate 1,000 Hz). Skin conductance (SC) data were collected with two standard Ag/AgCl electrodes attached to the second and third distal phalanges of the participant's non-dominant hand. Electric shocks were delivered transcutaneously through the participant's fourth and fifth distal phalanges of a dominant hand using standard Ag/AgCl electrodes.

Experimental procedure and Design

Prior to the experiment, an information email was sent to all participants including the description of the study, the exclusion criteria, and Covid-related information applied at the time of the experiment. Participants were informed that the experiment will take place in the lab and that they are going to make a number of decisions together with other participants. They were further informed that their participation will be anonymous and that their identity will not be revealed to other participants. Although participants were in the same room, their anonymity was protected through the following measures: (i) they were asked to arrive at the lab with a 5-minute difference, so they will not meet each other before, during, and after the experiment, (ii) they were seated in separate cubicles and they were not able to see each other during the experiment, and iii) the participants were asked to wear ear mufflers during the experiment, so they will not hear each other (e.g. keyboard presses, reactions to shock administration).

Upon arrival, participants were asked to read and sign the consent form. Subsequently, heart rate (HR), skin conductance, and shock electrodes were attached and informed by the researcher that they can start the experiment.

Participants first completed the following questionnaires: the State and Trait Anxiety Inventory (STAI, Spielberger et al., 1970), to measure anxiety level, the Social Value Orientation task (SVO, Murphy et al., 2011) to measure social preferences, and the Staircase Risk Elicitation Task (RET, (Falk et al., 2016; Holzmeister, 2017) to measure risk preferences (**Figure 2**). In the current study, only the Social Value Orientation task was analysed.

During the SVO task, the participants were asked to make fifteen decisions on how to allocate money between themselves and the other participants in the lab. In each decision, they could either allocate the money in a way that maximises their own pay-off (self-serving, i.e., 100 points allocated to oneself and 50 points to the other person) or in a way that benefits the other person (altruistically, e.g., 50 points allocated to one-self and 100 points to the other person. The pattern of their decisions determined their "social value orientation angle", with larger the values indicating more baseline prosociality (Murphy et al., 2011).

Next, the shock calibration procedure was performed individually (see above) and participants were asked to read the instructions for the public good game and answer comprehension questions. Participants had to answer all the comprehension questions correctly before starting the public goods game.

Upon completion of the public good game, participants completed demographic questions (age, gender, and field of study), and debriefing questionnaires e.g., about whether they knew other participants in the study or how many shocks they have received. Finally, participants were debriefed and paid. During debriefing, they were informed about the purpose of the experiment. The total earnings were the sum of the Social Value Orientation Task (Murphy et al., 2011), the Staircase Risk Elicitation Task

(Falk et al., 2016), and one block of the public goods game. The block of the public goods game that was used was chosen randomly.

Public goods game

To examine our hypotheses of the effect of threat and punishment on cooperation we used a linear public goods game (Wit & Wilke, 1998). The task was performed in a group of 3 participants. Participants performed the public good game under the threat condition, which was the within-subjects factor, with two levels ("safe" = no electric shocks were received, and "threat" = electric shocks) and under punishment conditions, which was the between-subjects factor, with, also, two levels ("peer-punishment" and "control"). Threat and safe conditions were indicated through a red and green colour of the participants' icons, respectively, displayed on screen during the decision and feedback parts of a trial. When the icons were red, participants could receive an electric shock to their fingers. Shocks were incidental, i.e., independent of behavioural responses, and all participants received them simultaneously. No shocks were delivered in safe trials. Feedback followed by an ITI (inter-trial interval) of 2-4 seconds.

At the beginning of each trial, each participant received an endowment of 10 MU. On each trial, participants had to decide how much of this endowment they wanted to contribute to the group pool (reflecting other-regarding preferences) and how much to keep for themselves (reflecting self-regarding preferences). A total contribution to the group pool was multiplied by a factor of 1.5 and divided equally among the three participants. The public goods game poses a social dilemma to groups. A total benefit of

15 MU is achieved for the members of the group, when all group members fully cooperate, by investing all of their resources to the public good. While the full contribution of all group members to the group pool maximizes group earnings, individual group members earn even more by withholding their own MU and benefit from the contributions of others (free-riding). For example, if one member contributes zero (0) (free-riding) to the group pool and the other two members fully contribute (10 x 2 x 1.5= 30 MUs in the group pool), the free-rider will gain 20 MUs (10 MUs, his own endowment + 10 MUs from others contribution to the common pool), which are more than the 15 MUs that this individual would earn if fully contributed to the common pool. After the decision phase, feedback displayed for 10 seconds containing information on individual and others' contributions to the group pool and everyone's earnings from a given trial. The duration of the decision phase was 10 seconds as this period was the minimum duration required for the acquisition of physiological data. If participants did not make a decision within the decision window, their contribution was chosen randomly by the computer, but the participants themselves earned 0 in that specific trial. This was done to encourage participants to make active decisions on each trial. This procedure was known to the participants beforehand.

Additionally, half of the tested groups (30 groups) were exposed to a peer punishment procedure. Specifically, following the feedback phase, participants had an option to punish other participants by means of deducting Deduction Points (DPs) from other participants' accounts. The cost of each deducted DP was 1 MU for the decision-maker, and 4 MU for the punished participant. After the deduction phase, participants received feedback on how many DPs each of them had deducted and received in total (they did not know from whom they received the DPs exactly). These two phases of peer punishment (decision making and feedback) were not time-constrained.

In the current study, a 2 x 2 mixed factorial design was conducted. We used threat condition as a within-subject condition with two levels: safe and threat, and peer punishment as a between-subject condition with two levels: presence and absence of punishment. Each condition was organised in blocks of 15 consecutive trials. Threat condition was counterbalanced, i.e., half of the groups started with a safe, and another half with a threat condition. Shock trials represented 10% of all trials, i.e. 3 trials with shocks during the whole task. All tested groups (60 groups) were exposed to both levels of threat condition (i.e. threat and safe). Additionally, half of the tested groups (30 groups) were exposed to a peer punishment procedure.

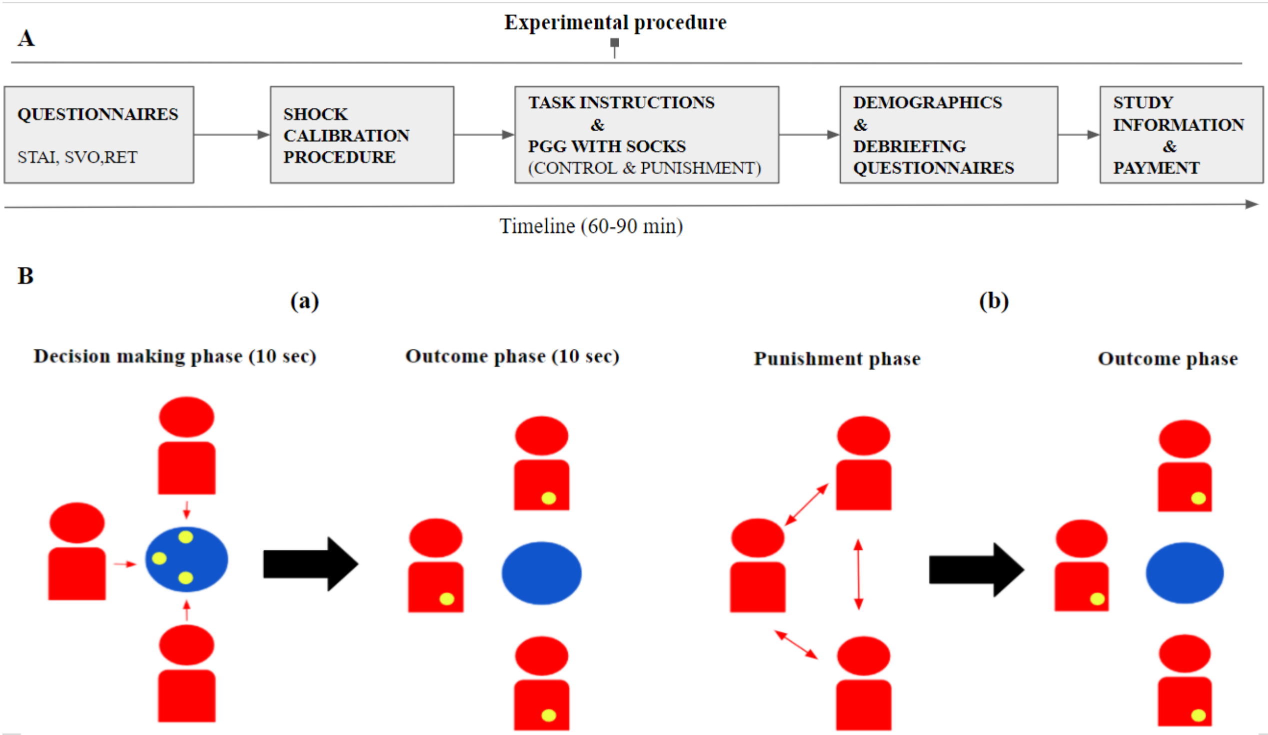

Figure 2. Experimental procedure and the public goods game. A) Experimental procedure; STAI, State and Trait Inventory; RET, Staircase Risk Elicitation Task; SVO, Social Value Orientation, PGG; Public Goods Game. The duration of the experiment was about 60-90 min. B) Public goods game. (a) Control (no-punishment) condition. In the decision making phase, participants (in a group of 3, red icons) decided on how much of their initial endowment to contribute to a public good (blue circle), keeping the remaining amount. In the outcome phase (10sec), participants were informed of their individual contributions to the group pool and their earnings. (b) Punishment condition. Half of the tested groups were additionally exposed to a peer punishment procedure. During the punishment phase, participants had an option to punish other participants by means of deducting Deduction Points (DPs) from other participants' accounts. In the outcome phase, participants were informed on how many DPs each of them had deducted and received in total.

Data Analysis

The statistical analyses of the behavioural and physiological data were performed in R (R version 4.1.2: 2021-11-01) and SPSS (IBM SPSS Statistics for Windows, Version 27.00). The analyses were performed using Repeated Measures ANOVA, using ezANOVA function (ez package, version 4.4.0), and regression analysis, using lm function (DAAG package, version1.24). In order to examine whether our threat manipulation was successful, we assessed the effect of threat on heart rate (HR, reflecting parasympathetic activity), and skin conductance levels (SCL, reflecting sympathetic activity) during the public goods game. Baseline-corrected HR and SCL responses (i.e., means of HR during the decision making phase were subtracted from the baseline represented by the mean of the HR in 20-sec breaks preceding a given block of trials were calculated for each participant. A repeated-measures ANOVA was conducted with the threat condition as within-subjects factor, the punishment conditions as a between-subjects factor, and the baseline-corrected HR responses as the dependent variable. The same analysis was performed for SCLs.

Individual contributions to the public good in the public goods game (range 0-10 MU) were used as an index of cooperation. In order to measure the effect of threat and punishment on cooperation, a repeated-measures ANOVA with threat condition (safe, threat) as within-subjects factor, peer punishment (present, absent) as the between-subjects factor, and individual mean contributions as the dependent variable were conducted.

Additionally, we were also interested to what extent the threat and punishment manipulations influence free-riding behaviour in the public goods game. Free-riding was observed in those trials where participants contributed nothing (0 MU) to the group pool. For our analyses, we calculated the percentage of free-riding decisions participants made in each condition. In order to measure the effect of threat and punishment conditions on free-riding, a repeated-measures ANOVA with threat condition (safe, threat) as within-subjects factor, peer punishment (present, absent) as a between-subjects factor, and the percentage of individual free-riding decisions as the dependent variable was conducted.

Finally, we tested whether the number of assigned deduction points in punishment conditions differs between threat and safe trials. To this end, a repeated-measures ANOVA with threat condition (safe, threat) as within-subjects factor and the mean of assigned deduction points as the dependent variable was conducted.

It should be noted that since the participants performed the task in groups, behavioural responses between these participants may be interdependent. Because it is not feasible to conduct nested analysis in ANOVA, in our analyses, we did not account for the fact that participants were nested in groups.

Results

Physiological results

In order to determine whether our threat manipulation was successful, we examined the effect of threat conditions on HR and SCL. We first tested the effect of threat on heart rate reflecting parasympathetic activation. The results showed that threat resulted in changes in heart rate (HR), $F(1,142) = 4.88$, $p = .028$, $\eta^2 = .01$, with lower HR in the threat ($M = -2.43$, $SD = 7.05$) compared to safe condition ($M = -1.25$, $SD = 6.42$), **(Figure 3)**. HR did not differ between control and punishment conditions, as indicated by a non-significant main effect of punishment $F(1,142) = 3.23$, $p = .074$, $\eta^2 = .01$. Finally, the interaction between threat and punishment conditions was non-significant, $F(1,142)=1.28$, $p = .25$, $\eta^2 = .003$ suggesting that the effect of threat on HR did not differ between control and punishment conditions.

Next, we tested the effect of threat on SCL reflecting sympathetic activation. We found that threat induced larger SCL responses $F(1,142)= 65.06$, $p < .001$ $\eta^2 = .20$, with higher SCL in threat ($M = 1.19$, $SD = 2.61$) compared to safe condition ($M = -0.68$, $SD = 1.75$), **(Figure 3)**. Neither punishment, $F(1,142)=0.28$, p=.59, $\eta^2 < .001$ nor its interaction with threat condition $F(1,142)=1.64$, $p=.20$, $\eta^2 = .006$ were significant.

These results support the conclusion that our threat manipulation was successful in evoking a typical physiological response pattern under threat, and that these responses did not differ as a function of punishment condition.

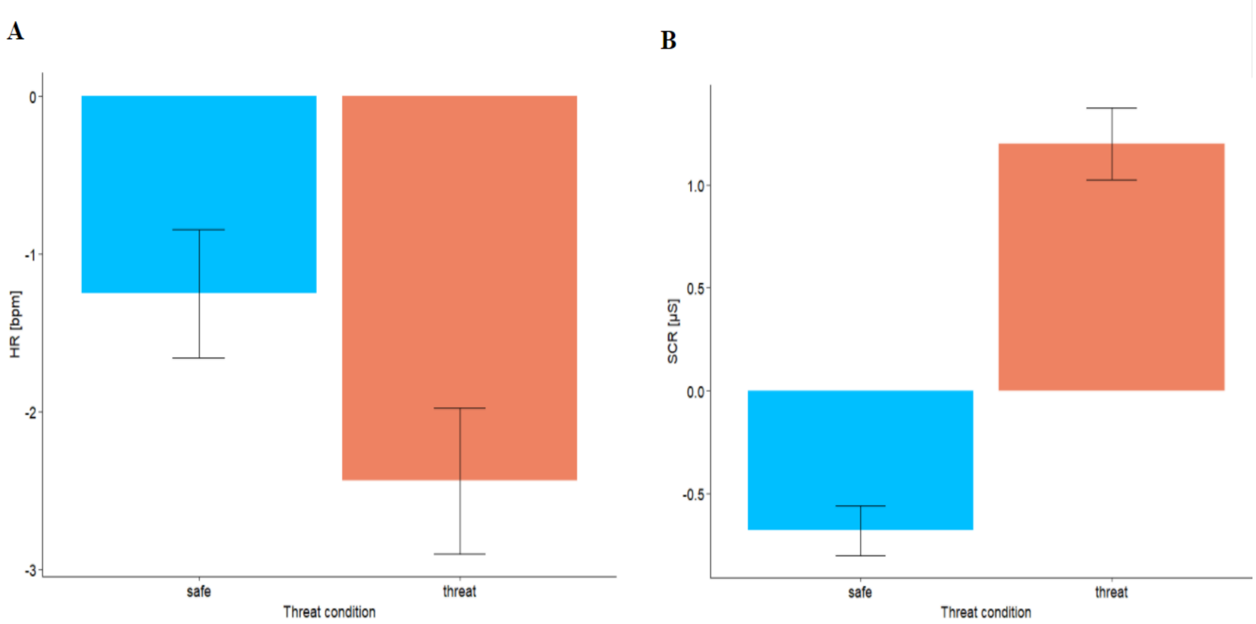

Figure 3. Physiological responses during the public goods game. (A) Baseline-corrected heart rate (HR) responses. Lower heart rate (HR) responses in threat compared to safe conditions reflect parasympathetic activation. (B) Baseline-corrected skin conductance levels (SCL) responses. Higher skin conductance levels (SCL) in threat compared to safe condition reflect sympathetic activation. µS: microSiemens; bpm: beats per minute. Error bars represent standard errors of the mean.

Behavioural results

We first tested the effect of threat and punishment conditions on cooperation. Cooperation was defined by participants' contributions (in MU) to the group pool in the public goods game. We found a significant main effect of the punishment condition on cooperation, $F(1,178)=15.15$, $p <.001$, $\eta^2 = 0.07$ with higher overall contributions in punishment ($M = 6.82$, $SD = 2.05$) compared to control condition ($M = 5.51$, $SD = 2.65$) **(Figure 4)**. This result is in line with the previous studies (Fehr & Schurtenberger, 2018; Lohse & Waichman, 2020; Molleman et al., 2019) where punishment was successful in increasing overall contributions to the public good. The main effect of threat $F(1,178)=0.10$, $p = .74$, $\eta^2 < .001$ was not significant, suggesting that cooperation did not differ between safe and threat conditions. Finally, the interaction between threat and punishment conditions was not significant $F(1,178)=0.41$, $p=.51$, $\eta^2 < .001$, suggesting that the effect of threat on cooperation did not differ between control and punishment conditions.

Next, we examined the effect of threat and punishment conditions on free-riding. Free-riding was defined as 0 MU contributions to the group pool in the public goods game. First, we found a significant main effect of punishment on free-riding $F(1,178)=8.89$, $p=.003$, $\eta^2= 0.03$ with more free-riding decisions observed in control ($M = 6.60$, $SD = 2.83$) compared to punishment condition ($M = 2.05$, $SD = 1.59$) (**Figure 5)**, further confirming that our punishment manipulation was successful. The main effect of threat $F(1,178)=0.26$, $p=.61$, $\eta^2< .001$, was not significant suggesting that free-riding was not different in safe compared to threat condition. Finally, the interaction between threat and punishment conditions $F(1,178)=1.04$, $p=.30$, $\eta^2 = .001$, in free-riding was also not significant suggesting that the effect of threat on free-riding was not different in control compared to punishment conditions.

In our next analysis, we focused on the deduction points and whether the magnitude of assigned deduction points differed between threat and safe trials. Deduction points were defined by the number of Deduction Points (DP) participants had deducted from other participants' accounts. The analysis revealed that there was no difference between threat and safe conditions in the number of Deducted Points during punishment condition, $F(1,89)=2.99$, $p=.08$, $\eta^2 = .002$.

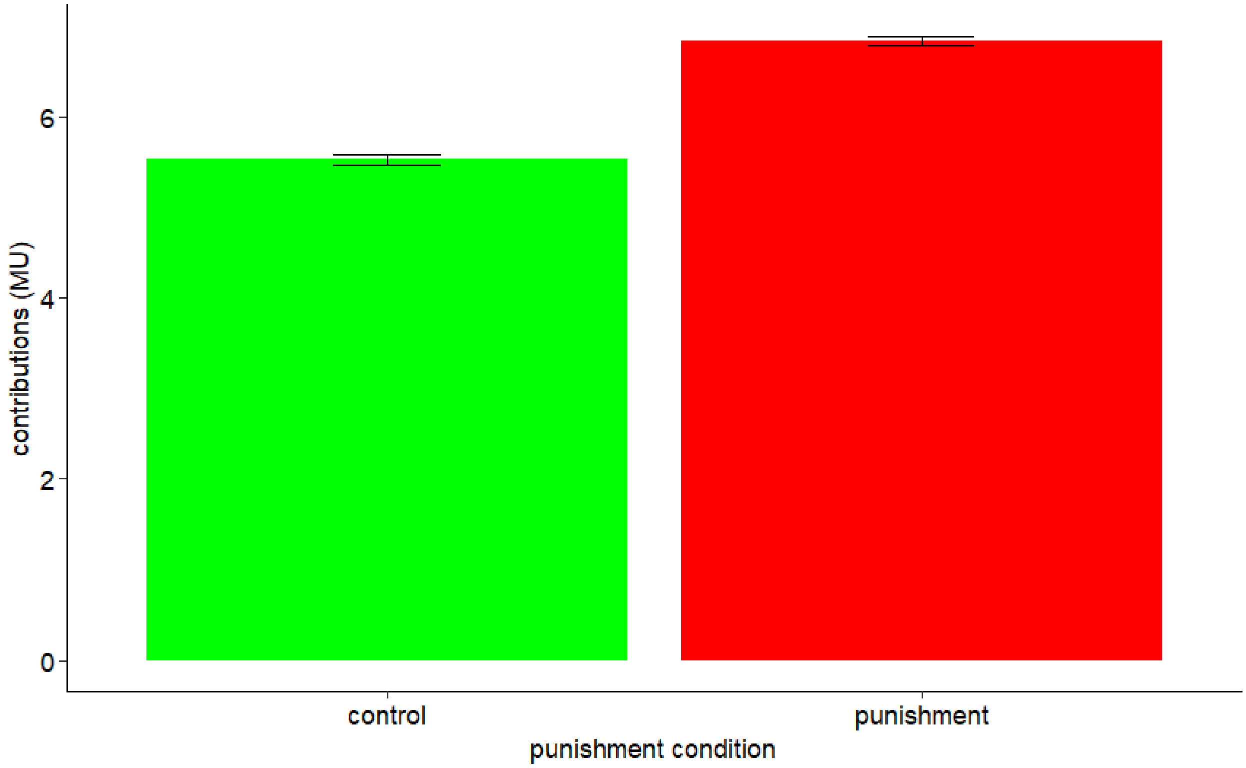

Figure 4. Mean contributions to the public good. Participants contributed more in punishment compared to control condition. MU, Monetary Units

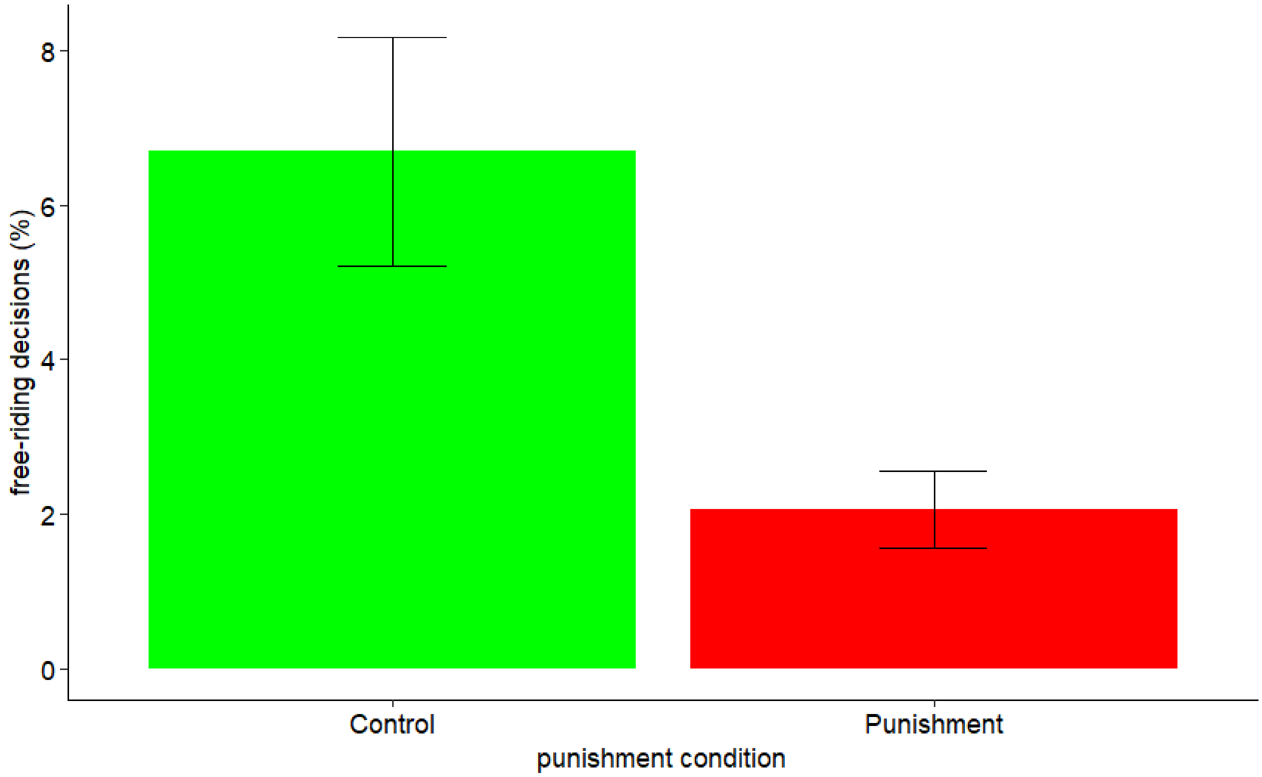

Figure 5. Fee-riding decisions in the public goods game. Free-riding decisions are represented by null contributions to the group pool. Higher percentage of free-riding decisions in control compared to punishment condition

Exploratory analyses

The relationship between SVO and contributions

Previous studies have shown that individual baseline prosociality modulated the relationship between threat exposure and cooperation, with higher baseline prosociality associated with more cooperation under threat (Lojowska et al., 2022). Thus, in the next exploratory step, we aimed to investigate whether and how baseline prosociality moderates the effect of threat and punishment on cooperation. To this end, we performed repeated-measures ANOVA, with the threat condition as within-subjects factor, the punishment conditions as a between-subjects factor, the SVO angle as a covariate, and mean individual contributions as the dependent variable. The analyses revealed a significant main effect of baseline prosociality, $F(1,176) = 6.09$, $p = .015$, $\eta^2 = .033$ with larger baseline prosociality associated with more cooperation **(Figure 6)**. The interactions between baseline prosociality and punishment condition $F(1,176) = 2.00$, $p = .158$, $\eta^2 = .011$, as well as between baseline prosociality and threat conditions, $F(1,176) = .359$, $p = .550$, $\eta^2 = .002$, were non-significant. Finally, the interaction between threat condition, punishment condition, and baseline prosociality as also non-significant, $F(1,176) = .059$, $p = .808$, $\eta^2 < .001$. These results imply that whereas more prosocial individuals contributed overall more to the public good, this relationship did not differ between threat and punishment conditions.

The relation between physiological threat responses and contributions

There are findings supporting that the physiological responses to threat influence decision-making (Fanselow, 1994; Öhman et al., 2001). For example, Lojowska and colleagues (2022) found that reduced HR responses to threat exposure were associated with smaller contributions to the group pool, particularly among less prosocial individuals. Thus, in our next exploratory step, we examined the relationship between a physiological index of freezing, i.e., reduction in heart rate and its association with contribution decisions. To this end, we performed a regression analysis with the difference in heart rate between threat versus safe condition, and punishment condition as the independent variables, and the difference in contributions between threat and safe conditions as the dependent variable. First we found that threat-related changes in heart rate did not predict contributions in threat vs. safe condition, $B < -.001$, $t(140) = .01$, $p = .986$. This relationship also did not differ between control and punishment conditions, as supported by a non-significant interaction between threat-related changes in HR and punishment condition $B = .005$, $t(140) = .15$, $p = .87$.

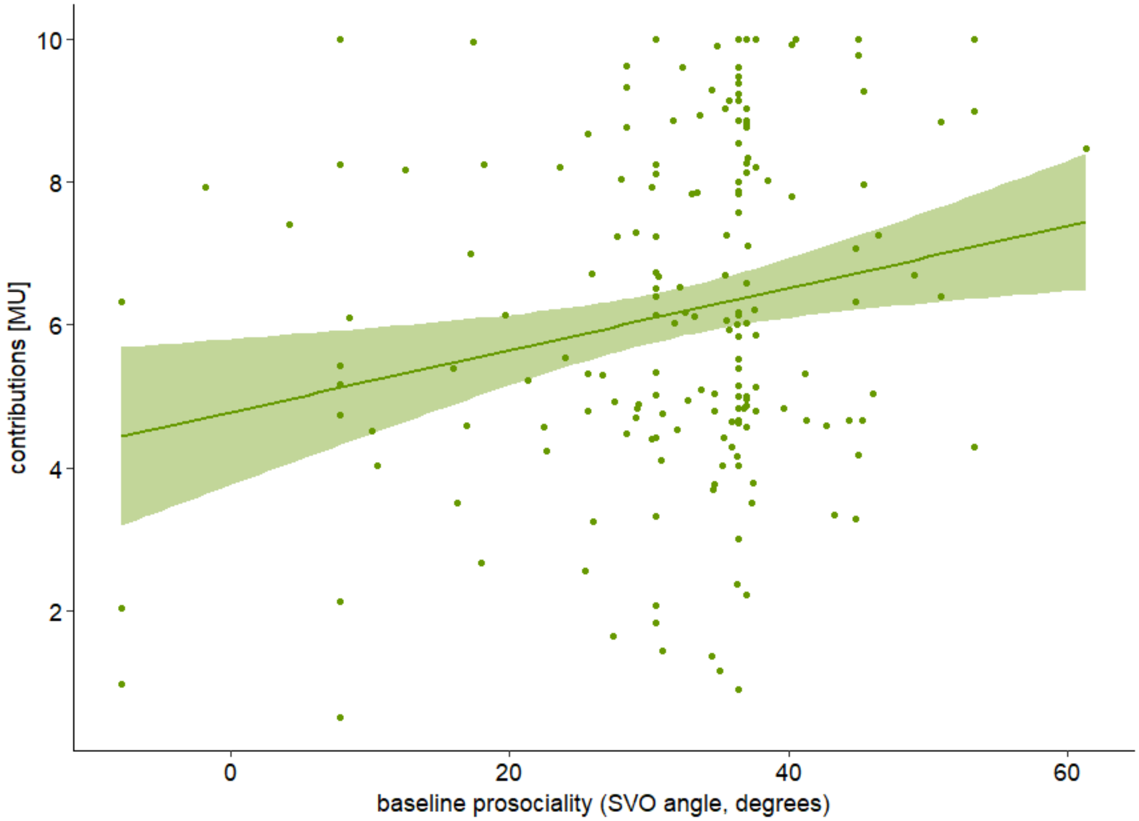

Figure 6. Contributions to the public good and baseline prosociality. Baseline prosociality is represented by SVO angle with a higher SVO angle representing higher prosociality. Higher contributions indicate more cooperation. Higher prosociality is related to more contributions regardless of the effect of threat and punishment.

Discussion

Human groups are exposed to environmental and economic threats that can profoundly affect individual survival and group functioning. There is evidence suggesting that the threatening conditions facilitate affiliation (Mawson, 2005; Mobbs et al., 2020) and avert the collapse of cooperation which is normally observed in non-threatening situations (Lojowska et al., 2022). In the current study, we examined whether in the presence of threat, less punishment is used to maintain cooperation. We found that participants in our experiment did not exhibit more cooperation under threat, as supported by similar contributions to the group pool under threat and in the safe condition. In line with the previous studies (Fehr & Fischbacher, 2003; Fehr & Gächter, 2002), we also found that peer-punishment increased overall cooperation, supported by higher levels of cooperation when the option of punishment was given compared to control (no-punishment) condition. Here, we were primarily interested in whether, provided a facilitating effect of threat on cooperation, cooperation would be maintained with less punishment. However, we found that people employed punishment to the same degree under threat and in the safe condition, as indicated by the similar number of deduction points received in these two conditions. Thus, we can conclude that punishment increased cooperation overall, and higher levels of cooperation under threat were not associated with less punishment i.e., assigned deduction points to others in the group. Consequently, the presence of threat does not diminish the role of punishment in increasing cooperation.

There is research to support that the experience of threat can promote cooperation (Lojowska et al., 2022). Surprisingly, our results indicate that people express similar levels of cooperation regardless of whether they are facing a threat or not. Importantly, our results are not possible due to unsuccessful threat manipulation in our experiment. It has been found, that at the physiological level, defensive responses to threat manifest as sympathetically controlled increases in skin conductance and parasympathetically driven decreases in heart rate (Gladwin et al., 2016; Mackersie & Calderon-Moultrie, 2016; Roelofs, 2017). Based on this evidence, in the current study, we recorded participants' heart rate (HR) and skin conductance levels (SCL) and the results reveal that under threat, participants expressed a typical freezing reaction with lower heart rate and higher skin conductance. Thus, we assume that our threat manipulation is successful. It is important to note, however, that previous research found the effect of threat on cooperation over time. Specifically, it is known that the cooperation among people, although high at the beginning, deteriorates over time due to the conflict between self-interest and collective interest (Fehr & Gachter, 2000; van Dijk & De Dreu, 2021). Threat exposure has been proposed as a means to slow down the decline in cooperation over time (Lojowska et al., 2022). However, in the current study, we measured the main effect of threat on cooperation rather than an overtime effect, suggesting that although threat exposure did not affect participants' overall cooperation, the facilitating effect of threat could be still observed over time. Another factor that could explain why threat exposure does not affect cooperation is the order in which participants were exposed to threat and safe conditions. In our study, the order of threat conditions was randomized, with half of the groups starting with a safe condition, and the other half with a threat condition. It is possible that a facilitating effect of threat in cooperation could be observed when threat condition presented first, and, hence, there was no prior interaction compared to when threat presented second. That is because in safe conditions, a reduced pattern of cooperation is expected over time. Thus, the expected reduction of cooperation over time, could override the effect of threat in cooperation, when the experiment started with the safe condition. To identify a potential order effect in the current study, we could have analysed the main effect of order in cooperation i.e., the difference in the levels of cooperation between the groups who assigned to the threat condition first compared to the groups who assigned to the threat condition second. One solution in order to control a potential order effect would be the duration between threat and safe condition to be increased. In conclusion, our study supports that threat exposure does not facilitate cooperation. Some reasons that might explain this lack of effect, is that firstly, we measured the overall effect of threat in cooperation and not the effect across trials, and secondly, that the order in which threat and safe condition were presented to the participants might have affected their contributions

There is evidence that threat-related changes in physiology influence human cooperation (Lojowska et al., 2022). Since in the current study, here participants expressed a typical freezing reaction under threat, too, we could expect that it would influence their behaviour. However, we did not find an association between individual physiological responses to threat, i.e. heart rate reduction, and cooperation. One potential reason why we do not see threat-related physiological responses to predict cooperation is that previous research has shown that the association between the physiological defensive responses and cooperation is observed specifically among less prosocial individuals (Lojowska et al., 2022). In our study, prosociality found to predict cooperation, with higher prosociality to be associated with more cooperation. However, this effect was not different between threat and safe condition. Nevertheless, human prosociality could modulate the effect of threat-related physiological activation in cooperation. In conclusion, our study supports that the physiological activation under threat does not facilitate cooperation. Future studies could investigate whether prosociality influences the effect of individual defensive responses to overall cooperation.

We found that punishment increased cooperation, which was evident in general contributions to the group pool, and reduced free-riding. Our findings on the impact of peer punishment on human cooperation are consistent with previous research linking peer punishment to improved cooperation and limited free-riding (Fehr & Gächter, 2002; Lohse & Waichman, 2020; Nikiforakis, 2010). Because of previous research suggesting a threat-facilitating effect on cooperation, we hypothesize that group members do not need to invoke peer punishment to maintain cooperation. Indeed, when the possibility of punishment was offered, participants achieved similar and importantly high levels of cooperation in both safe and threat condition. However, in contrast to our expectations, cooperation under threat is not achieved with less punishment, as it would have been expected based on a previously found facilitating effect of threat on cooperation. One of the possible explanation could be that we do not see the effect of threat in the first place. Since the effect of threat is not found to promote cooperation in this study, the high cooperation in the event of punishment could be due exclusively to the facilitating effect of punishment on cooperation, rather than the interaction between threat and punishment. Another explanation could be that threat does not facilitate cooperation overall, and that is the reason why a similar level of punishment is required to facilitate cooperation on threat and safe condition. Overall, our study suggests that the presence of threat does not diminish the role of punishment in increasing cooperation. We showed that threat by itself does not promote cooperation and peer punishment is needed for increasing cooperation in threatening situations.

Theoretical and practical implications

The results of the current study allow us to observe the results of previous studies on decision-making under threat from a different perspective. Although previous research has linked threat exposure and peer punishment with higher cooperation (Lojowska et al., 2022; Fehr & Gächter, 2002), to our knowledge, we are the first to investigate the necessity of punishment in threatening situations. In social decision-making settings, the results of the current study could illuminate whether costly regulations are required to increase cooperation in threatening conditions. By implication, they inform whether and when humans should organize institutions for the maintenance of cooperation, and when such institutions are not needed. Contemporary societies seek to discover strategies to promote cooperation in the face of pressing threats such as environmental and health crises. Our results show that threat by itself does not promote cooperation and that peer punishment is needed for increasing cooperation in threatening situations. For instance, those results help us understand why cooperation breaks down during COVID-19 pandemic (Talor & Habibi, 2020). Some people act as free-riders and neglect the safety measures that are imposed by the government. Despite not contributing to or even compromising public safety, those people benefit from others' adherence to policies, which reduces their health risks (Yong & Choy, 2021). Our results imply that the threat imposed by being infected with coronavirus is not adequate to increase cooperation, and that is the reason why some people do not comply with the measures. In this case, based on our results we can conclude that peer punishment was required to enforce cooperation during the COVID-19 pandemic. Overall, our study imply that societies could benefit from developing strategies that enable peer punishment within the community to promote cooperation in threatening situations.

Limitations and future directions

Although the present study expands the literature on cooperation under threat and punishment its findings should be considered within its methodological limitations. Since we conducted a laboratory experiment there were elements that we controlled and could hinder results' transferability. Real-life situations are characterised by inequality among individual resources. For example, some people are wealthier than others or individual solutions to shared problems are more easily accessible to them. There is evidence that inequality in resources could result in the collapse of cooperation (Burton-Chellew et al., 2013; Marotzke et al., 2020). Additionally, Gross and Dreu (2019) showed that when people have the option to solve a collective problem individually, they are less willing to cooperate with others. Importantly, they suggest that the breakdown of cooperation is amplified instead of mitigated by peer punishment. However, in our experiment, all participants received the same endowment. Thus, we were unable to examine whether the asymmetry which occurs in real-life situations affects the degree to which punishment is required to preserve cooperation under threat. Since punishment is costly, we could expect that people who are able to be preserved in threatening situations by individual solutions, might employ punishment less compared to safe situations. That is because under threat, people need to preserve their personal resources in order to invest them in individual solutions. However, in safe conditions, people do not need to invest their resources in order to be preserved, thus, they might be more willing to punish others in order to maintain cooperation. Future studies could incorporate this aspect of asymmetry, and instead of providing the same endowment to all participants divide participants into "rich" and "poor". Thus, it could be examined whether the interaction between cooperation and punishment under threat is influenced by inequality.

Another factor that could influence the transferability of the results of the current study is the way that threat was presented. Specifically, in our study, threat was instrumental to cooperation, meaning that participants did not have the option to avoid threat. However, in real life, usually, threat is instrumental, implying that people have the option to avoid an aversive situation if they cooperate. The possibility of threat avoidance through cooperation could influence the levels of punishment required for cooperation to be maintained. For example, it has been found that people living in harsh environments exhibit high levels of cooperation without punishment (Lewis et al., 2014). In our study, the fact that participants did not have the option to avoid punishment could decrease their willingness to cooperate, increasing the need for punishment in order for the cooperation to be maintained in the group. However, if participants could cooperate to avoid threat, higher levels of cooperation could be observed without the need for peer punishment.

Future studies could investigate this relationship by giving the option to participants to cooperate and achieve some joined target in order to avoid losing their endowment.

Conclusions

Overall, our study highlights that threat by itself does not promote cooperation and peer punishment is needed to sustain cooperation in threatening situations. Although previous research has linked punishment with increased cooperation, to our knowledge, our study is the first to investigate the necessity of punishment in promoting cooperation under threat. In today's society, identifying tactics that enable cooperation is critical since humans are continuously exposed to a range of threats, varying from environmental and economic shocks to public health risks, that can severely impact individual survival and group functioning. Based on this study, societies could develop strategies that enable peer punishment within the community to promote cooperation in threatening situations. However, our results should be interpreted with caution given that they might be caused by the lack of the facilitating effect of threat in human cooperation in the first place. We suggest that threat does not promote cooperation in our study because we measured the overall effect of threat in cooperation, whereas previous research supports that threat maintains cooperation over time. Future research could investigate whether punishment is necessary for promoting cooperation under threat when threat is instrumental and when individual solutions to shared problems is an option.

References

Balafoutas, L., Nikiforakis, N., & Rockenbach, B. (2014). Direct and indirect punishment among strangers in the field. *Proceedings of the National Academy of Sciences*, *111*(45), 15924–15927. https://doi.org/10.1073/pnas.1413170111.

Blanchard, D. C., Griebel, G., Pobbe, R., & Blanchard, R. J. (2011). Risk assessment as an evolved threat detection and analysis process. *Neuroscience & Biobehavioral Reviews*, *35*(4), 991–998. https://doi.org/10.1016/j.neubiorev.2010.10.016

Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3281–3288. https://doi.org/10.1098/rstb.2009.0134.

Burton-Chellew, M. N., May, R. M., & West, S. A. (2013). Combined inequality in wealth and risk leads to disaster in the climate change game. *Climatic Change*, *120*(4), 815–830. https://doi.org/10.1007/s10584-013-0856-7.

Calo-Blanco, A., Kovářík, J., Mengel, F., & Romero, J. G. (2017). Natural disasters and indicators of social cohesion. *PLOS ONE*, *12*(6), e0176885. https://doi.org/10.1371/journal.pone.0176885

Change, *10*(6), 518–525. https://doi.org/10.1038/s41558-020-0783-3

Dal Bo, P. (2005). Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. *American Economic Review*, *95*(5), 1591–1604.

de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., & al, et. (2004). The Neural Basis of Altruistic Punishment. *Science*, *305*(5688), 1254–1258.

Dezecache, G., Grèzes, J., & Dahl, C. D. (2017). The nature and distribution of affiliative behaviour during exposure to mild threat. *Royal Society Open Science*, *4*(8), 170265.

Dezecache, G., Martin, J.-R., Tessier, C., Safra, L., Pitron, V., Nuss, P., & Grèzes, J. (2021). Nature and determinants of social actions during a mass shooting. *PLOS ONE*, *16*(12), e0260392. https://doi.org/10.1371/journal.pone.0260392

Dietz, T., Ostrom, E., & Stern, P. C. (2003). The Struggle to Govern the Commons. *Science*. https://doi.org/10.1126/science.1091015

Engel, C. (2014). Social preferences can make imperfect sanctions work: Evidence from a public good experiment. *Journal of Economic Behavior & Organization*, *108*, 343–353. https://doi.org/10.1016/j.jebo.2014.02.015

Engelmann, J. B., Meyer, F., Ruff, C. C., & Fehr, E. (2019). The neural circuitry of affect-induced distortions of trust. Science Advances, 5, eaau3413.

Falk, A., Becker, A., Dohmen, T. J., Huffman, D., & Sunde, U. (2016). *The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences* (SSRN Scholarly Paper ID 2725874). Social Science Research Network. https://doi.org/10.2139/ssrn.2725874

Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin & Review*, *1*(4), 429–438.

Fanselow, M. S., & Lester, L. S. (1988). A functional behavioristic approach to aversively motivated behavior: Predatory imminence as a determinant of the topography of defensive behavior. In *Evolution and learning* (pp. 185–212). Lawrence Erlbaum Associates, Inc.

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*(2), 175–191. https://doi.org/10.3758/BF03193146

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*(6960), 785–791. https://doi.org/10.1038/nature02043

Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, *8*(4), 185–190. https://doi.org/10.1016/j.tics.2004.02.007

Fehr, E., & Gachter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review, 90*(4), 980–994.

Fehr, E., & Gächter, S. (2001). *Do Incentive Contracts Crowd Out Voluntary Cooperation?* (SSRN Scholarly Paper ID 289680). Social Science Research Network. https://papers.ssrn.com/abstract=289680

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140.

Fehr, E., & Schmidt, K. M. (2010). On inequity aversion: A reply to Binmore and Shaked. *Journal of Economic Behavior & Organization*, *73*(1), 101–108. https://doi.org/10.1016/j.jebo.2009.12.001

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour*, *2*(7), 458–468. https://doi.org/10.1038/s41562-018-0385-5

FeldmanHall, O., Raio, C. M., Kubota, J. T., Seiler, M. G., & Phelps, E. A. (2015). The Effects of Social Context and Acute Stress on Decision Making Under

Uncertainty. *Psychological Science*, *26*(12), 1918–1926. https://doi.org/10.1177/0956797615605807

Fowler, J. H. (2005). Altruistic punishment and the origin of cooperation. *Proceedings of the National Academy of Sciences*, *102*(19), 7047–7049.

g, P. J., Simons, R. F., Balaban, M., & Simons, R. (2013). *Attention and Orienting: Sensory and Motivational Processes*. Psychology Press.

Gladwin, T. E., Hashemi, M. M., van Ast, V., & Roelofs, K. (2016). Ready and waiting: Freezing as active action preparation under threat. *Neuroscience Letters*, *619*, 182–188.

Gross, J., & Dreu, C. K. W. D. (2019). Individual solutions to shared problems create a modern tragedy of the commons. *Science Advances*. https://doi.org/10.1126/sciadv.aau7296

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, *7*(1), 1–16. https://doi.org/10.1016/0022-5193(64)90038-4

Hamilton, W. D. (1971). Geometry for the selfish herd. *Journal of Theoretical Biology*, *31*(2), 295–311. https://doi.org/10.1016/0022-5193(71)90189-5

Hardin, G. (2009). The Tragedy of the Commons. *Journal of Natural Resources Policy Research*, *1*(3), 243–253. https://doi.org/10.1080/19390450903037302

Hauert, C., Traulsen, A., Brandt, H., Nowak, M. A., & Sigmund, K. (2007, December). *The Emergence of Altruistic Punishment: Via Freedom to Enforcement* [Monograph]. IR-07-053. http://pure.iiasa.ac.at/id/eprint/8405/

Hauser, O. P., Rand, D. G., Peysakhovich, A., & Nowak, M. A. (2014). Cooperating with the future. Nature, 511(7508), 220–223. https://doi.org/10.1038/nature13530

Holzmeister, F. (2017). oTree: Ready-made apps for risk preference elicitation methods. *Journal of Behavioral and Experimental Finance*, *16*, 33–38. https://doi.org/10.1016/j.jbef.2017.08.003

Ioannou, C. (2021). Grouping and Predation. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 3574–3580). Springer International Publishing. https://doi.org/10.1007/978-3-319-19650-3_2699

Kallhoff, A. (2014). Why societies need public goods. *Critical Review of International Social and Political Philosophy*, *17*(6), 635–651. https://doi.org/10.1080/13698230.2014.904539

LanEngel, C. (2014). Social preferences can make imperfect sanctions work: Evidence from a public good experiment. *Journal of Economic Behavior & Organization*, *108*, 343–353. https://doi.org/10.1016/j.jebo.2014.02.015

Lehtonen, J., & Jaatinen, K. (2016). Safety in numbers: The dilution effect and other drivers of group life in the face of danger. *Behavioral Ecology and Sociobiology*, *70*(4), 449–458. https://doi.org/10.1007/s00265-016-2075-5

Leimar, O., & Hammerstein, P. (2001). Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1468), 745–753. https://doi.org/10.1098/rspb.2000.1573

Lewis, H. M., Vinicius, L., Strods, J., Mace, R., & Migliano, A. B. (2014). High mobility explains demand sharing and enforced cooperation in egalitarian hunter-gatherers. *Nature Communications*, *5*(1), 5789. https://doi.org/10.1038/ncomms6789

Li, X., Molleman, L., & van Dolder, D. (2021). Do descriptive social norms drive peer punishment? Conditional punishment strategies and their impact on cooperation. *Evolution and Human Behavior*, *42*(5), 469–479. https://doi.org/10.1016/j.evolhumbehav.2021.04.002

Lohse, J., & Waichman, I. (2020). The effects of contemporaneous peer punishment on cooperation with the future. *Nature Communications*, *11*(1), 1815. https://doi.org/10.1038/s41467-020-15661-7

Lojowska, M. Gross, J. De Dreu, C.K.W. (2022) *Threat prevents the breakdown of cooperation* [Unpublished manuscript]. Institute of Psychology, Leiden University.

Lojowska, M., Ling, S., Roelofs, K., & Hermans, E. J. (2018). Visuocortical changes during a freezing-like state in humans. *NeuroImage*, *179*, 313–325. https://doi.org/10.1016/j.neuroimage.2018.06.013

Löw, A., Weymar, M., & Hamm, A. O. (2015). When threat is near, get out of here: Dynamics of defensive behavior during freezing and active avoidance. *Psychological Science*, *26*(11), 1706–1716.

Mackersie, C. L., & Calderon-Moultrie, N. (2016). Autonomic nervous system reactivity during speech repetition tasks: Heart rate variability and skin conductance. *Ear and Hearing*, *37*, 118S-125S.

Marotzke, J., Semmann, D., & Milinski, M. (2020). The economic interaction between climate change mitigation, climate migration and poverty. *Nature Climate*

Mawson, A. R. (2005). Understanding mass panic and other collective responses to threat and disaster. *Psychiatry: Interpersonal and Biological Processes*, *68*(2), 95–113.

Milinski, M., Semmann, D., Krambeck, H.-J., & Marotzke, J. (2006). Stabilizing the Earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences*, *103*(11), 3994–3998. https://doi.org/10.1073/pnas.0504902103

Milinski, M., Sommerfeld, R. D., Krambeck, H.-J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, *105*(7), 2291–2294. https://doi.org/10.1073/pnas.0709546105

Mobbs, D., Headley, D. B., Ding, W., & Dayan, P. (2020). Space, Time, and Fear: Survival Computations along Defensive Circuits. *Trends in Cognitive Sciences*, *24*(3), 228–241. https://doi.org/10.1016/j.tics.2019.12.016

Molina, J. A., Giménez-Nadal, J. I., Cuesta, J. A., Gracia-Lazaro, C., Moreno, Y., & Sanchez, A. (2013). Gender Differences in Cooperation: Experimental Evidence on High School Students. *PLOS ONE*, *8*(12), e83700. https://doi.org/10.1371/journal.pone.0083700

Molleman, L., Kölle, F., Starmer, C., & Gächter, S. (2019). People prefer coordinated punishment in cooperative interactions. *Nature Human Behaviour*, *3*(11), 1145–1153. https://doi.org/10.1038/s41562-019-0707-2

Nikiforakis, N. (2010). Feedback, punishment and cooperation in public good experiments. *Games and Economic Behavior*, *68*(2), 689–702. https://doi.org/10.1016/j.geb.2009.09.004

Nowak, M. A., & Sigmund, K. (1998). The Dynamics of Indirect Reciprocity. *Journal of Theoretical Biology*, *194*(4), 561–574. https://doi.org/10.1006/jtbi.1998.0775

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466.

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413–425. https://doi.org/10.1016/j.tics.2013.06.003

Roelofs, K. (2017). Freeze for action: Neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160206.

Shentu, T., Ma, J., & Guo, Y. (2018). Social attachment shapes emergency responses: Evidence from a postfire study. *Social Behavior and Personality: An International Journal*, *46*(1), 139–150.

Sokolov, E. N., & Cacioppo, J. T. (1997). Orienting and defense reflexes: Vector coding the cardiac response. In *Attention and orienting: Sensory and motivational processes* (pp. 1–22). Lawrence Erlbaum Associates Publishers.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). Stai. *Manual for the State-Trait Anxiety Inventory (Self Evaluation Questionnaire). Palo Alto California: Consulting Psychologist*, *22*, 1-24.

Taylor, A. L., & Habibi, R. (2020). The collapse of global cooperation under the WHO International Health Regulations at the outset of COVID-19: sculpting the future of global health governance. ASIL Insights, 24(15), 15.

Tedeschi, E., Buyalskaya, A., & Silston, B. (2020). Subjective Fear Dilutes as Group Size Increases.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/10.1086/406755 van Dijk, E., & De Dreu, C. K. W. (2021). Experimental Games and Social Decision Making. *Annual Review of Psychology*, *72*(1), 415–438. https://doi.org/10.1146/annurev-psych-081420-110718

Van Lange, P. A. M., & Rand, D. G. (2022). Human Cooperation and the Crises of Climate Change, COVID-19, and Misinformation. *Annual Review of Psychology*, *73*(1), 379–402. https://doi.org/10.1146/annurev-psych-020821-110044

Vieira, J. B., Pierzchajlo, S., Jangard, S., Marsh, A., & Olsson, A. (2020). *Perceived threat and acute anxiety predict increased everyday altruism during the COVID-19 pandemic*.

Vugt, M. V., Cremer, D. D., & Janssen, D. P. (2007). Gender Differences in Cooperation and Competition: The Male-Warrior Hypothesis. *Psychological Science*, *18*(1), 19–23. https://doi.org/10.1111/j.1467-9280.2007.01842.x

Wang, X., Ding, R., Zhao, J., Chen, W., & Gu, C. (2021). *The Effect of Inequity Aversion Driven Punishment on Cooperation* (SSRN Scholarly Paper ID 3996845). Social Science Research Network. https://doi.org/10.2139/ssrn.3996845

Wit, A., & Wilke, H. (1998). *Public good provision under environmental and social uncertainty*. https://doi.org/10.1002/(SICI)1099-0992(199803/04)28:2<249::AID-EJSP868>3.0.CO;2-J

Yahaghi, R., Ahmadizade, S., Fotuhi, R., Taherkhani, E., Ranjbaran, M., Buchali, Z., Jafari, R., Zamani, N., Shahbazkhania, A., Simiari, H., Rahmani, J., Yazdi, N., Alijani, H., Poorzolfaghar, L., Rajabi, F., Lin, C.-Y., Broström, A., Griffiths, M. D., & Pakpour, A. H. (2021). Fear of COVID-19 and Perceived COVID-19 Infectability Supplement Theory of Planned Behavior to Explain Iranians'

Intention to Get COVID-19 Vaccinated. *Vaccines*, *9*(7), 684. https://doi.org/10.3390/vaccines9070684

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and social Psychology*, *51*(1), 110.

Yong, J. C., & Choy, B. K. C. (2021). Noncompliance With Safety Guidelines as a Free-Riding Strategy: An Evolutionary Game-Theoretic Approach to Cooperation During the COVID-19 Pandemic. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/article/10.3389/fpsyg.2021.646892