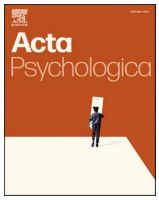
Acta Psychologica 262 (2026) 106119

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy



Emotion matters: The pathways and the consequence of third-party intervention triggered by unfairness perception

Jiachen Qu [a], Jan Drewes [a], Yunfeng He [b], Jiajin Yuan [a,*]

[a] The Affect Cognition and Regulation Laboratory (ACRLab), Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu, China
[b] Liaoning Key Laboratory of Psychological Testing and Behavior Analysis, Liaoning University, 66 Chongshan Road, 110036, Shenyang, China

ARTICLE INFO

Keywords:
Third-party intervention
Altruistic behavior
Emotion
Injustice perception
Compensation

ABSTRACT

Third-party intervention is an altruistic behavior of maintaining fairness at personal costs in a self-irrelevant situation. Though prosocial emotion has been considered an important trigger of altruistic behavior, how emotional involvement is related to third-party intervention during unfair events remains elusive. To address this issue, we manipulated allocation fairness and assessed emotional states before and after intervention in a third-party dictator game. Participants were assigned different intervention abilities and given three choices: punishment, compensation, or inaction. The results indicate that unfair events trigger third-party intervention by eliciting injustice-related negative emotions. Increased event unfairness led to more frequent and intense intervention, along with a preference for punishment over compensation. Moreover, choosing to intervene resulted in a significant emotional improvement compared to inaction. Individuals with stronger intervention abilities or a heightened perception of unfairness experienced greater emotional improvement following intervention. Furthermore, compensatory intervention led to greater emotional improvement than punitive intervention. In summary, the present study demonstrates a mediating role of negative emotion in the impact of unfair events on third-party intervention which, in turn, alleviates the intensity of negative emotions associated with injustice perception.

1. Introduction

Third-party intervention (TPI) has been widely used to assess individuals' willingness to sacrifice self-interest and maintain fairness for a violation that is not directly self-related, which is crucial for upholding social norms justice (Dhaliwal et al., 2021; Fehr & Fischbacher, 2004a, 2004b; Gummerum et al., 2016; Leliveld et al., 2012; Van Doorn et al., 2018). However, the reasons why individuals are willing to engage in self-sacrifice without obtaining material benefits remain elusive. Researchers regard TPI as an evolutionary form of altruistic behavior, which supports long-term social stability and development (Fehr & Fischbacher, 2004a, 2004b; Krasnow et al., 2016). It has been suggested that intervention helps individuals build a personal reputation or establish connections and cooperation with others—two instrumental explanations for TPI in specific situations (Dhaliwal et al., 2021; Gummerum et al., 2016; Van Doorn et al., 2018). Nevertheless, the general and direct trigger for self-sacrifice remains unclear and requires further exploration. Accordingly, several studies indicate that individuals experienced third-party emotions such as anger and compassion during witnessing moral violations or unfair events (Ginther et al., 2022; Landmann & Hess, 2017). Evidence has shown that experiencing prosocial emotions, such as compassion and gratitude, is linked to intervention behaviors like third-party punishment in transgression situations (Pfattheicher et al., 2019; Vayness et al., 2020). Recent research has further demonstrated that the severity of the transgression modulates third-party emotional responses and intervention. Specifically, when a perpetrator actively 'harms others' rather than 'not helping others,' third parties report stronger moral anger and are more likely to punish (Yang, Gu, et al., 2022). This finding highlights that both the characteristics of the event and the evoked moral anger play a critical role in triggering TPI, which hints at the possibility that emotional arousal serves as a psychological bridge between social motivation and intervention choice. Despite these findings, the specific role of emotion evoked by unfair events in the decision to intervene and the choice of intervention methods has yet to be directly investigated (Landmann & Hess, 2017; O'Reilly et al., 2016).

The connection between emotion and intervention may not be unidirectional. Apart from the initial emotional impulse, the outcome of the

* Corresponding author at: Institute of Brain and Psychological Sciences, Sichuan Normal University, China.
E-mail address: yuanjiajin168@126.com (J. Yuan).

https://doi.org/10.1016/j.actpsy.2025.106119
Received 20 April 2025; Received in revised form 20 November 2025; Accepted 9 December 2025
Available online 19 December 2025

intervention may in turn regulate individual's feelings. Barasch et al. (2014) proposed that emotional benefit is a robust drive to altruistic behavior, where individuals may engage in such behavior to repair negative emotions. This contrasts with other instrumental explanations of prosocial behavior, such as the desire to build a reputation (Berman et al., 2015; Jordan et al., 2016). Meta-analytical findings from 201 studies revealed a robust connection between prosocial behavior and human well-being, encompassing hedonic well-being, psychological functioning, and physical health (Hui et al., 2020). Moreover, research indicates that following descriptive social norms can help individuals regulate their emotions and reduce negative feelings (Bergquist & Ekelund, 2025). Therefore, when a critical social norm like fairness is violated, it triggers intense negative affective responses in third parties, such as other-focused empathy and self-focused distress (Quoidbach et al., 2015). Perhaps these emotions directly prompt the third party to intervene. In turn, individuals also experience emotional gratification from their benevolent actions (Fig. 1).

Different unfair events trigger distinct emotions and influence the willingness to intervene. However, behavior still varies across individuals in the face of the same event, influenced by their intervention abilities (Dhaliwal et al., 2021; Ginther et al., 2022; Landmann & Hess, 2017; Leliveld et al., 2012; Van Doorn et al., 2018). In third-party intervention studies, the intervention ability is typically manipulated in the intervention cost-impact ratio, which reflects an individual's level of self-sacrifice in achieving the same intervention impact (Cheng et al., 2022). Intervention ability represents an individual's ability to exert control over and effectively change the outcome of an unfair event, which also reflects an individual's perceived sense of power to manipulate the situation. For instance, armed police exhibit a superior intervention ability compared to ordinary citizens in violent crime scenarios, which stems from their access to critical resources unavailable to citizens. The armed police have state-sanctioned authority, which provides them with legitimacy and effective coercive instruments, such as weaponry and the legal mandate to use it. Conversely, citizens lacking both official legitimacy and effective coercive means have a limited repertoire of intervention options and may incur personal risk.

Studies have shown that improved intervention ability could increase the willingness of individuals to punish (Cheng et al., 2022; Qu et al., 2018). Enhanced intervention ability provides individuals with a strong resource, and they can effectively use strategies to achieve their goals. Empirical studies have shown that individuals with high power tend to experience more positive emotions compared to those with low power (Berdahl & Martorana, 2006). These positive emotions, in turn, enhance self-expression related to active aspects of the self, such as self-regulation and prioritizing efforts toward focal goals (Guinote, 2017). In this perspective, stronger intervention ability not only facilitates engagement in intervention behaviors but also strengthens the emotional benefits of such actions. Nevertheless, no study has yet explored the impact of intervention ability on an individual's emotional state during TPI.

Clarifying emotional changes is valuable for understanding the choice between different TPI behaviors, such as punishment and compensation. Participants typically have two intervention options: punishing the perpetrator or compensating the victim. Evidence suggests that, despite both being methods of third-party intervention, punishment and compensation are not equally perceived from the observer's perspective. Compensation tends to convey more positive reputational signals, as acting without considering personal costs

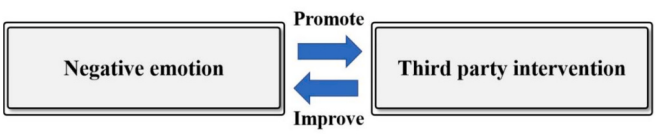


Fig. 1. The research theoretical hypothesis framework diagram.

increases perceived trustworthiness in helpers but not in punishers (Engeler & Raihani, 2024). This finding highlights a critical asymmetry in how altruistic versus punitive interventions are socially interpreted. Research indicates a general preference for compensation over punishment during TPI, with the choice of compensation being significantly higher than that of punishment (Dhaliwal et al., 2021; Van Doorn et al., 2018; Van Doorn & Brouwers, 2020). Nevertheless, this tendency may vary depending on the third party's focus or the severity of the transgression. Evidence shows that shifting the focus of the third party can alter intervention patterns, with a focus on the perpetrator leading to more punishment, while a focus on the victim increases the likelihood of compensation (David et al., 2017). Moreover, it has been postulated that greater transgression severity may shift intervention preferences from compensation to punishment, as fairness is a core social norm in decision-making (Van Doorn & Brouwers, 2020; Weiß et al., 2023). However, no research has yet explored the relationship between emotions and different intervention choices. We hypothesize that more unfair events may lead to heightened negative emotions, which in turn trigger a greater preference for punishment, due to anger-induced focus on the perpetrator (David et al., 2017). We also aim to explore how the chosen intervention method, in turn, influences emotional change. Punishment as a means of upholding fairness works by diminishing the perpetrator's benefits, serving as a form of retaliation against unfairness. In contrast, compensation may be a more effective approach for emotional restoration, as it helps elevate the victim from a disadvantaged position (Van Doorn & Brouwers, 2020). We predict that a preference for compensation should be associated with more efficient emotional rehabilitation. However, when the unfairness of an event becomes extreme, individuals are likely to choose punishment as an intervention due to anger directed at the perpetrator.

In summary, the present study investigates: 1) the influence of event unfairness and intervention ability on TPI behavior, and 2) the role of emotion in the occurrence of these effects. We measured third party's emotional states in a distribution scenario and manipulated event unfairness, while intervention ability is operationalized as the ratio of intervention influence to payment. Our main hypotheses are as follows: (1) increased event unfairness is associated with a higher rate of intervention and a greater preference for punishment, especially under conditions of higher intervention ability; (2) emotion mediates the effect of event unfairness on TPI behavior. Specifically, negative emotions induced by unfair events trigger intervention behavior, with heightened negative emotions driving individuals to punish the perpetrator; (3) the intervention has benefits in improving emotions. Specifically, stronger intervention ability enhances this effect, and compensation leads to greater emotional improvement compared to punishment.

2. Methods

2.1. Participants

We performed a sample size estimate using G*Power 3.1.9.7 (Faul et al., 2009). Based on a medium effect size (d = 0.25), an expected power value (1 – β = 0.80), and a significance level (α = 0.05), at least 93 participants were required. Considering potential sample loss, 120 healthy right-handed participants, aged 17 to 23 years, were recruited for the experiment and randomly assigned to three groups with different intervention abilities. All participants completed the experiment in the laboratory. However, fifteen participants were excluded from analysis: three misunderstood the roles of players A and B across trials, three missed the intervention choice, and nine did not take on the role of player C throughout the experiment. Consequently, 105 participants (67 females) with 35 in each group were included in the analysis. Gender distribution ($\chi^2 = 0.330$, $p = .848$) and age ($F(2,102) = 0.265$, $p = .768$, $\eta^2 = 0.005$) did not differ across the three groups.

All procedures in this study were conducted in accordance with the principles outlined in the Declaration of Helsinki 2004 and its

subsequent amendments, or comparable ethical standards. The experiment was approved by the local Ethics Committee of Sichuan Normal University. All participants provided written informed consent prior to their participation in the experiment.

2.2. Experimental design and procedure

We employed a modified Third-Party Dictator Game (TP-DG) based on Fehr and Fischbacher (2004b). In the standard TP-DG, three players are involved: Player A is asked to allocate a certain amount of Monetary Units (MUs) between themselves and Player B, with Player B having no choice but to accept the allocation. A third party, Player C, who is unrelated to the allocation, then decides whether to punish Player A by using their MUs to reduce A's payoff. In our study, we introduced a compensation option, allowing Player C to add to Player B's payoff (Gummerum et al., 2016). Punishment and compensation could not be selected simultaneously in a round of intervention. To make the amounts more suitable for purchasing power in China, Player A split 50 MUs (50 yuan), and Player C started with 25 MUs (25 yuan) in this experiment (Cheng et al., 2022).

Participants were informed that they were part of the second segment of a series of experiments. In the first segment, hundreds of participants were divided into groups, with two individuals in each group collaborating to complete a task. After the task, one person (Player A) was randomly selected to distribute the entire reward from the experiment (50 MUs, equivalent to 50 yuan), while the other person (Player B) could only accept the allocation. Player A had three allocation options: (1) A receives 25, B receives 25 (fair condition), (2) A receives 35, B receives 15 (moderately unfair condition), or (3) A receives 45, B receives 5 (highly unfair condition). Participants were informed that the allocations of Player A and B would be distributed after the second segment.

The participants observed the allocations of 30 different groups from task 1, with the order of the three allocations randomized by a computer program. Participants acting as Player C, could choose to intervene by either punishing Player A (reducing A's amounts) or compensating Player B (adding to B's amounts), with intervention costing Player C's MUs. For each round of intervention, Player C had 25 MUs and could spend up to 10 MUs on the intervention. The upper limit for the intervention value was set to simulate real-life situations, where individuals typically do not sacrifice all their interests when helping others. We categorize intervention ability into three conditions: Strong - where intervention significantly impacts the event at low cost, Medium - where intervention cost is equal to impact, and Weak - where impact is smaller than cost. The participants' intervention ability was randomly assigned before the experiment. Strong, medium, and weak intervention ability was defined by each MU spent for the intervention being converted to 3, 1, or 0.8 MUs subtracted from A's (punitive) or added to B's (compensatory) allocation result, respectively. For example, participants in the strong intervention ability group were told 'You will hold 25 MUs and can intervene in the above distribution results by spending your own amount. Every 1 MU spent can punish A or compensate B 3 MUs'. Additionally, all participants were informed that if they chose to intervene, the rewards for A and B would be determined by the results of their intervention; otherwise, the rewards would be based on A's allocation.[1]

Therefore, the experiment followed a 3 (event unfairness: fair, moderately unfair, highly unfair; within-subject variable) × 3 (intervention ability: weak, medium, strong; between-subject variable) mixed design. The experimental procedure consisted of three sections: the questionnaire section, the practice section, and the intervention section. Given that potential individual factors might influence intervention (FeldmanHall et al., 2015; Hu et al., 2015; Van Doorn & Brouwers, 2020), all participants were required to complete two questionnaires before experiment to confirm their empathy and baseline emotional feelings: (1) Interpersonal Reactivity Index (IRI) (Davis, 1980), which consists of 22 items to assess empathy across four dimensions: perspective-taking (PT), fantasy (FA), empathetic concern (EC), and personal distress (PD). Participants rated each item on a scale from 1 to 5 (1 = highly disagree, 5 = highly agree). The Cronbach's α in this study was 0.774. (2) The Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988). The schedule consists of 20 different emotional words (10 positive and 10 negative), and participants were asked to rate their emotions on a scale from 1 to 5 (1 = not at all, 5 = extremely). Reliability was acceptable for both the positive (α = 0.807) and the negative (α = 0.898) words. The Chinese version of both scales was used in this study, which has been tested for reliability and validity (Huang et al., 2003; Zhang et al., 2010).

After filling out the questionnaires, participants proceeded to a practice section where they practiced using a slider to select options. This included a practice segment identical to the formal intervention, consisting of 3 trials, to familiarize participants with the intervention procedure. For both rating and intervention sliders, participants can use '1' in the numpad to move left and '3' to move right. The intervention section consisted of 30 trials: 10 fair, 10 moderately unfair, and 10 highly unfair, presented randomly (Fig. 2a). In each trial, participants viewed an allocation result for 3 s, then they were required to rate the perceived injustice of the allocations and their emotional state at the moment (Injustice perception: How just do you think this allocation is? 1 extremely unjust ~9 extremely just; Emotional state: Your current emotional feelings? 1 extremely unpleasant ~9 extremely pleasant, pre-decision emotional state rating). Subsequently, participants had to choose one from punishment, compensation, or inaction by pressing 'f', 'j', or 'b' on the keyboard, respectively. If they chose to punish or compensate, they had up to 30 s to determine the amount spent for intervention and could confirm their decision by pressing the '0' on the keyboard numpad. For those choosing inaction, a blank screen appeared randomly for 2 to 6 s (Fig. 2b). All participants rated their current emotional state again after the intervention page (post-decision emotional state rating).

2.3. Data analyses

Statistical analysis was performed using IBM SPSS Statistics software version 26.0 (IBM Corp., Armonk, NY). First, to confirm randomization, we conducted a one-way ANOVA on initial emotional state, empathy, and other factors to ensure there were no significant differences across the three groups. Subsequently, we used repeated measures ANOVA and PROCESS 4.1 (Preacher & Kelley, 2011) to explore the effects of event unfairness (fair, moderately unfair, highly unfair) on injustice perception and emotional state (pre-decision) as a manipulation check. A bootstrapping procedure with 5000 replications and a 95 % bias-corrected confidence interval (CI) was used to calculate all indirect effects in this study.

Secondly, we conducted a two-factor (event unfairness: fair, moderately unfair, highly unfair; intervention ability: weak, medium, strong) mixed ANOVA on intervention behavior, including the rate of intervention, the rate of punishment and that of compensation, along with a chi-square test to investigate the shift in intervention method preference across different unfair situations. Since Mauchly's test indicated that the assumption of sphericity was violated for event unfairness, the results were corrected using Greenhouse-Geisser estimates.

Moreover, we explored the role of emotional state as a mediator in the effect of event unfairness on intervention behavior, based on the observation of significant event unfairness effect on the rate of intervention and that of punishment. Event unfairness was treated as the predictor (0 = moderately unfair, 1 = highly unfair), emotional state (pre-decision) as the mediator, and the rate of intervention and that of

[1] The final reward of Player C is calculated as follows: 5 MUs basic reward + [(the MUs in the round with the most MUs left + the MUs in the round with the least MUs left)/2], so the final reward ranges between 20–30 MUs.

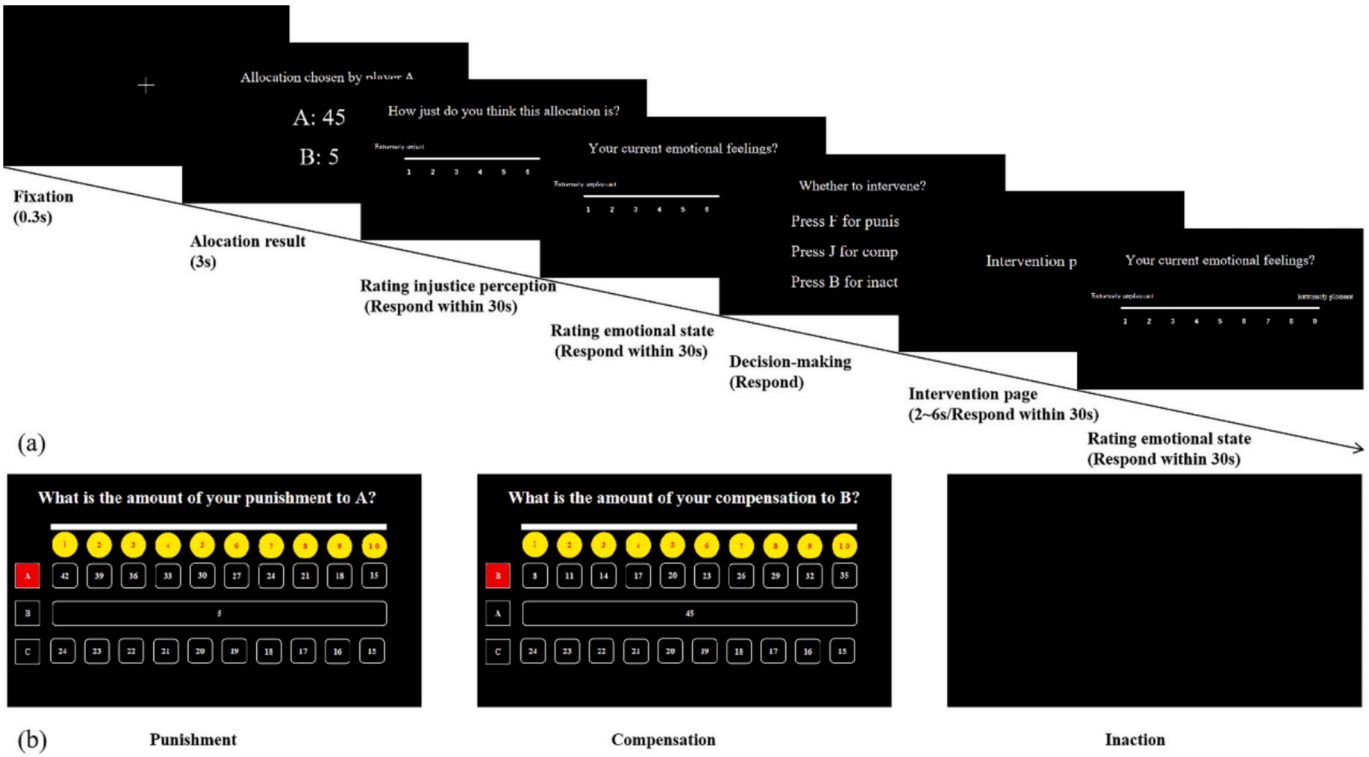


Fig. 2. (a) Third-party intervention task. (b) Examples of Intervention pages correspond to the highly unfair condition with strong intervention ability in decision-making.

punishment as the outcomes.

Thirdly, we examined how emotional state changes as a result of intervention behavior, and further explored how the emotional state change post versus pre intervention varied as a function of event fairness and intervention ability. Three steps of analysis were performed: 1) a two-factor (behavior: intervention, inaction; timepoint: pre-decision, post-decision) repeated measures ANOVA was conducted with emotional state as dependent variable. This analysis excluded data from participants who did not choose inaction during the experiment; 2) a three-factor (event unfairness: moderately unfair, highly unfair; intervention ability: weak, medium, strong; timepoint: pre-decision, post-decision) mixed ANOVA was conducted with emotional state as the dependent variable. This analysis used each participant's data when they chose to intervene in unfair events and excluded data from participants who didn't intervene in moderately or highly unfair events; 3) a two-factor (event unfairness: moderately or highly unfair; intervention method: punishment, compensation) ANOVA was conducted with emotional improvement as the dependent variable. This analysis only included data from participants who had used both intervention methods (punishment and compensation) throughout the experiment. Post hoc pairwise comparisons following significant main or interaction effects were corrected according to the Bonferroni method.

3. Results

3.1. Randomization check

Table 1 presents the results of randomization check, showing that there were no significant differences between the intervention groups in terms of age, IRI scores, and PANAS scores, as confirmed by the ANOVA analyses.

3.2. Manipulation check

The ANOVA was used to examine the effect of event unfairness on emotional state, with event unfairness (fair, moderately unfair, highly unfair) as independent variable, and emotional state (pre-decision) as dependent variable. Results showed that the effect of event unfairness on emotional state was significant, $F(2, 208) = 461.062$, $p < .001$, $\eta^2 = 0.816$. Further simple effect analysis revealed that the pre-decision

Table 1
Randomization check.

	Weak	Medium	Strong	$F_{(2,102)}$	Sig.
Age (years)	19.860 ± 1.683	19.860 ± 1.517	19.630 ± 1.330	0.161	ns
IRI	53.571 ± 7.972	54.886 ± 9.458	53.171 ± 8.648	0.691	ns
PT	11.714 ± 3.544	12.171 ± 2.935	11.371 ± 3.422	0.514	ns
FA	16.171 ± 2.792	16.314 ± 3.668	15.543 ± 3.657	0.511	ns
EC	16.086 ± 2.228	16.086 ± 3.442	16.714 ± 2.480	0.602	ns
PD	9.600 ± 4.202	10.314 ± 4.020	9.543 ± 3.567	0.417	ns
PANAS (P)	2.994 ± 0.582	3.006 ± 0.727	2.920 ± 0.720	0.164	ns
PANAS (N)	2.123 ± 0.683	2.303 ± 0.732	2.210 ± 0.710	0.560	ns

Note. (1) Weak: $N_{total} = 35$, $N_{female} = 23$; Medium: $N_{total} = 35$, $N_{female} = 23$; Strong: $N_{total} = 35$, $N_{female} = 21$. (2) Abbreviations: Sig., significance; ns, not significant; IRI, interpersonal reactivity index; PT, perspective-taking; FA, fantasy; EC, empathetic concern; PD, personal distress; PANAS, positive and negative affect schedule; P, positive; N, negative. (3) The data here are presented as Mean ± Standard Deviation.

emotional state became more negative as event unfairness increased (*ps* < 0.001).

We assessed potential mediation role of injustice perception in the effect of event unfairness on emotional states. Event unfairness as predictor, injustice perception as mediator, and pre-decision emotion as the outcome. The result showed a significant total effect (Total effect = −2.411, SE = 0.08, 95 % CI:[−2.57, −2.25]), while the direct effect of event unfairness was not significant (Direct effect = −0.108, SE = 0.15, 95 % CI:[−0.398, 0.181]). The mediating effect of event unfairness on emotion through injustice perception was significant (Indirect effect = −2.303, BootSE = 0.15, Boot95% CI:[−2.59, −2.02]) (Fig. 3). These data suggest that the current task effectively induced injustice perception which, in turn, elicited negative emotion.

3.3. The influence of event unfairness and intervention ability on intervention behavior

As described, a two-way mixed ANOVA was performed with event unfairness and intervention ability as independent variables, while the rate of intervention, that of punishment, and that of compensation as dependent variables.

The results for the rate of intervention revealed a significant main effect of event unfairness, $F(2,204) = 918.811$, $p < .001$, $\eta^2 = 0.900$, participants were inclined to intervene in more unfair events. However, there was no significant main effect of intervention ability, $F(2,102) = 0.976$, $p = .380$, $\eta^2 = 0.019$, suggesting that participants' intervention ability did not influence their willingness to intervene. No interaction effect was found, $F(4,204) = 0.801$, $p = .488$, $\eta^2 = 0.015$.

For the rate of punishment, we found a significant main effect of event unfairness, $F(2,204) = 150.893$, $p < .001$, $\eta^2 = 0.597$, with the rate of punishment increased during more unfair events. No significant effect of intervention ability ($F(2,102) = 0.594$, $p = .554$, $\eta^2 = 0.012$) or interaction effect ($F(4,204) = 0.407$, $p = .787$, $\eta^2 = 0.008$) was observed.

For the rate of compensation, we found a significant main effect of event unfairness, $F(2,204) = 120.871$, $p < .001$, $\eta^2 = 0.542$. The compensation rate was higher in moderately unfair and highly unfair compared to fair events ($p < .001$), while the two unfair conditions showed no significant differences ($p = .204$). No significant effect of intervention ability ($F(2,102) = 1.346$, $p = .265$, $\eta^2 = 0.026$) or interaction effect ($F(4,204) = 1.006$, $p = .399$, $\eta^2 = 0.019$) was observed. The response rates under different conditions were presented in Tables 2 and 3.

Based on the impact of event unfairness on intervention behavior, a Chi-square test indicated that individuals were more inclined to compensate the victim in moderately unfair situations, while they were more likely to punish the perpetrator in highly unfair situations ($\chi^2 = 45.838$, $p < .001$) (Table S1).

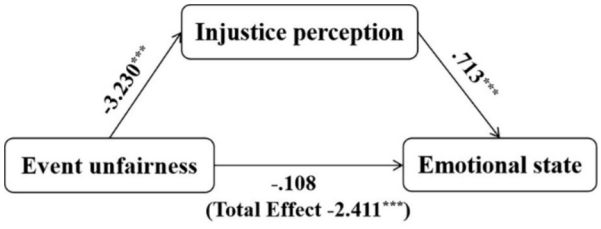


Fig. 3. Mediating effect of injustice perception in the effect of event unfairness on emotional state.

Note. (1) Unstandardized coefficients are displayed. (2) Event unfairness: 0 = fair, 1 = moderately unfair, 2 = highly unfair. Injustice perception: 1 = extremely injustice, 9 = extremely justice. Emotional state: 1 = extremely unpleasant, 9 = extremely pleasant. (3) *$p < .05$, **$p < .01$, and ***$p < .001$ (the same hereinafter).

Table 2
Response rates for different event unfairness ($N = 105$).

	Punishment (M ± SE)	M ± SE (compensation)	M ± SE (intervention)
Fair	0 ± 0	0.010 ± 0.003	0.010 ± 0.003
Moderately unfair	0.321 ± 0.027	0.477 ± 0.030	0.798 ± 0.028
Highly unfair	0.521 ± 0.029	0.407 ± 0.028	0.928 ± 0.014

Note. Abbreviations: M, mean; SE, standard error; N, sample size.

Table 3
Response rates for different intervention abilities ($N = 35$ for per group).

	M ± SE (punishment)	M ± SE (compensation)	M ± SE (intervention)
Weak	0.301 ± 0.025	0.265 ± 0.026	0.566 ± 0.021
Medium	0.262 ± 0.025	0.305 ± 0.026	0.567 ± 0.021
Strong	0.279 ± 0.025	0.324 ± 0.026	0.603 ± 0.021

Note. Abbreviations: M, mean; SE, standard error; N, sample size.

3.4. The mediation effect of emotional state

In unfair events (moderately or highly unfair), event unfairness influenced the rate of intervention and the choice of punishment, but not that of compensation. Therefore, we further assessed the mediating role of emotional state in the effect of event unfairness (moderately unfair, highly unfair) on the rate of intervention and punishment (outcome).

We first used the rate of intervention as the outcome. The result showed a significant total effect (Total effect = 0.130, SE = 0.03, 95 % CI:[0.068, 0.191]). However, after emotional state was included as a mediator, the direct effect of event unfairness (Direct effect = 0.072, SE = 0.04, 95 % CI:[−0.004, 0.147]) was not significant. Specifically, the mediating effect of emotional state subserving the effect of event unfairness on the rate of intervention was significant (Indirect effect = 0.058, BootSE = 0.025, Boot95% CI:[0.011, 0.110]) (Fig. 4a).

Moreover, we tested whether emotional state mediated the effect of event unfairness (predictor) on the rate of punishment (outcome). The result revealed a significant total effect (Total effect = 0.200, SE = 0.04, 95 % CI:[0.123, 0.278]), direct effect (Direct effect = 0.122, SE = 0.05, 95 % CI:[0.011, 0.028]), and indirect effect (Indirect effect = 0.078, BootSE = 0.03, Boot95% CI:[0.018, 0.135]) (Fig. 4b). In summary, event unfairness influenced participants' decision of intervention and the selection of punishment through emotional induction.

3.5. Emotional improvement effect of intervention

Firstly, we used a two-way ANOVA to investigate the emotional consequence of intervention behavior, with behavior (intervention, inaction) and timepoint (pre-decision, post-decision) as factors (Table S2). There was a significant main effect of timepoint, $F(1,74) = 883.458$, $p < .001$, $\eta^2 = 0.544$, and a significant interaction effect between behavior and timepoint, $F(1,74) = 15.629$, $p < .001$, $\eta^2 = 0.175$. Participants who chose to intervene exhibited a more negative emotional state than those who chose inaction in the pre-decision ($p < .001$), while this difference was absent in the post-decision ($p > .40$). Moreover, the post hoc paired samples *t*-test showed that a greater emotional improvement during intervention compared to inaction choices ($p < .001$) (Fig. 5).

Secondly, we tested how event fairness and intervention ability modulated emotional alterations following intervention decision, with event unfairness (moderately or highly unfair), intervention ability (weak, medium, strong), and timepoint (pre-decision, post-decision) as ANOVA factors (Table S3). The results showed a significant main effect of timepoint, $F(1,98) = 148.071$, $p < .001$, $\eta^2 = 0.602$, with the emotional state significantly more positive during post-decision versus

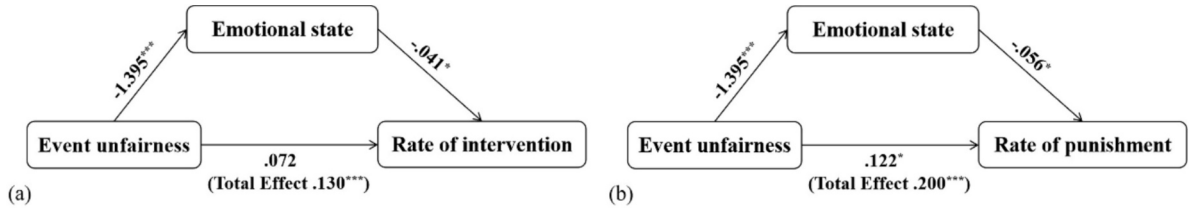


Fig. 4. The mediating effect of emotional state in the relationship between event unfairness and intervention behavior. *Note.* (1) Unstandardized coefficients are displayed. (2) Event unfairness: 0 = moderately unfair, 1 = highly unfair. Emotional state: 1 = extremely unpleasant, 9 = extremely pleasant.

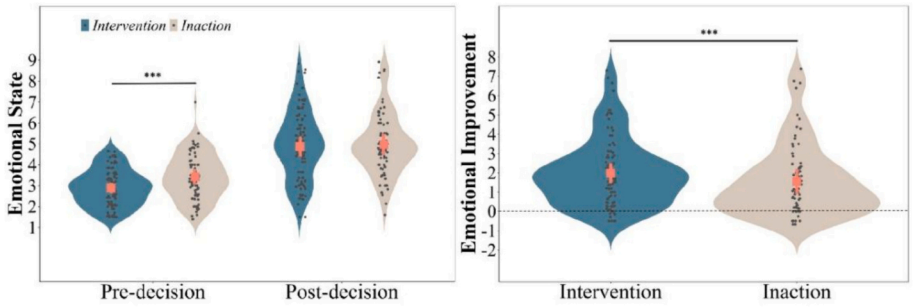


Fig. 5. Interaction effect between behavior and timepoint. *Note.* (1) Emotional improvement: pre-decision -post-decision in emotional state. Positive values indicate a positive shift in emotion, negative values indicate a negative shift, and zero indicates no shift. (2) Error bars represent standard errors. (3) Emotional state: 1 = extremely unpleasant, 9 = extremely pleasant.

pre-decision ($p < .001$). Additionally, there was a significant main effect of event unfairness, $F(1,98) = 263.146$, $p < .001$, $\eta^2 = 0.729$, and participants in highly unfair events exhibited more negative emotional states. The main effect of intervention ability was also significant, $F(2,98) = 6.519$, $p = .002$, $\eta^2 = 0.117$, and emotional state was overall more positive in strong compared to weak ($p < .001$) and medium ($p < .005$) ability groups, while the latter two groups showed no significant differences.

More importantly, there was a significant interaction effect of intervention ability × event unfairness, $F(2,98) = 3.144$, $p = .047$, $\eta^2 = 0.060$ (Fig. 6a). The strong intervention ability group showed significantly more positive emotional states than the weak group in both moderately ($p < .005$) and highly unfair conditions ($p < .001$), while the medium group didn't show significant differences from the other groups in both contexts. Moreover, there was a significant interaction effect of intervention ability × timepoint, $F(2,98) = 6.253$, $p = .003$, $\eta^2 = 0.113$ (Fig. 6b), with emotional states being significantly more positive in the strong compared to the medium ($p < .01$) and weak ($p < .001$) intervention ability groups after intervention choice. In contrast, the emotional state was not significantly different across the three groups

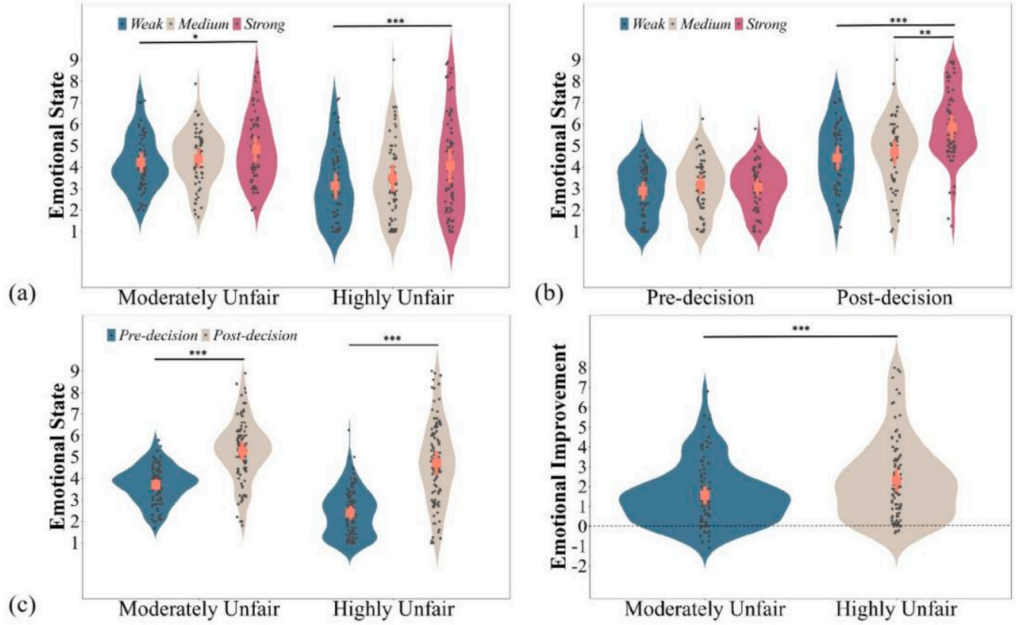


Fig. 6. The plots of emotional state measure varying as a function of (a) intervention ability × event unfairness interaction, (b) intervention ability × timepoint interaction, and (c) timepoint × event unfairness interaction. *Note.* (1) Emotional improvement: pre-decision - post-decision in emotional state. Positive values indicate a positive shift in emotion, negative values indicate a negative shift, and zero indicates no shift. (2) Error bars represent standard errors. (3) Emotional state: 1 = extremely unpleasant, 9 = extremely pleasant.

before the intervention choice. Lastly, there was an interaction effect of event unfairness × timepoint, $F(1,98) = 58.979$, $p < .001$, $\eta^2 = 0.376$ (Fig. 6c), indicating a significant improvement of emotional states after compared with before intervention choice for both moderately and highly unfair events ($p_s < 0.001$); however, highly unfair condition was associated with greater emotional improvement (Mean diff. = *2.32*) compared to moderately unfair condition (Mean diff. = *1.56*; $p < .001$).

Lastly, a two-way ANOVA was conducted to investigate potential interaction effects of event unfairness (moderately or highly unfair) and intervention method (punishment, compensation) on emotional improvement (post-decision - pre-decision) (Table S4). A significant main effect of the intervention method was found (Fig. 7), $F(1,66) = 5.651$, $p = .020$, $\eta^2 = 0.079$, showing that individuals opting for compensation experienced greater emotional improvement compared to those choosing punishment. However, no significant interaction effect was detected, $F(1,66) = 0.048$, $p = .827$, $\eta^2 = 0.001$, despite a main effect of event unfairness which confirmed greater emotional improvement in highly unfair events, $F(1,66) = 50.529$, $p < .001$, $\eta^2 = 0.434$.

4. Discussion

Prior studies indicate that perception of injustice or inequality is associated with decreased happiness and impaired well-being (Buttrick et al., 2017; Buttrick & Oishi, 2017), while prosocial emotions, such as other-focused anger or compassion, may trigger fairness-maintaining behavior like third-party punishment (Gummerum et al., 2016; Pfattheicher et al., 2019). Despite these hints, the role of injustice-related emotion in the influence of unfair events on third-party intervention has yet to be directly examined. The present study measured individuals' emotional states in pre- and post-decision in a variant of TPI paradigm, where perceived fairness and intervention ability were concurrently assessed. Our findings indicate that negative emotion triggered by unfair events mediates the effect of event fairness on TPI behavior which, in turn, alleviates the emotion intensity related to injustice perception.

4.1. Intervention selection increased with more unfair events but did not vary with intervention ability

We first analyzed intervention behavior across different levels of unfairness and intervention ability. Consistent with previous studies, both the rate of intervention and the choice of punishment were significantly influenced by the level of event unfairness (Cheng et al., 2022; David et al., 2017; Gummerum et al., 2016). This aligns with real-world behavior, where more severe or unjust events typically prompt more frequent intervention. Surprisingly, intervention abilities didn't affect the willingness to intervene, contrary to the findings of Cheng et al. (2022). This inconsistency may stem from differences in experimental design. In Cheng and colleagues' study, a within-subjects design was used, where participants' intervention ability (referred to as punishment cost) varied across trials. In contrast, the present study employed a between-subjects design, where participants were randomly assigned to different intervention ability conditions each, as manipulated, allowing participants a constant cost-effect ratio. Specifically, the within-subjects design in Cheng et al. (2022) encouraged participants to compare their intervention choices across trials, making them more sensitive to the changes in intervention abilities. This could have amplified the effect of intervention abilities on their decisions. However, in real-life scenarios, individuals typically have a stable influence over events. To simulate this situation, the current study adopted a between-subjects design to examine the influence of intervention ability. Accordingly, by assigning participants to a stable intervention ability, the effect of intervention ability on intervention decision could be performed and observed reliably, free of potential influence of trial-to-trial comparison.

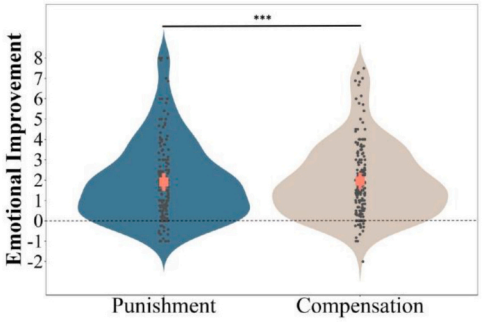


Fig. 7. The plot of the main effect of the intervention method. *Note.* (1) Emotional improvement: pre-decision - post-decision in emotional state. Positive values indicate a positive shift in emotion, negative values indicate a negative shift, and zero indicates no shift. (2) Error bars represent standard errors. (3) Emotional state: 1 = extremely unpleasant, 9 = extremely pleasant.

On the other hand, we observed an increased rate of intervention with more unfair events, even at the cost of personal income loss with the administration of third-party intervention decisions. This is consistent with prior findings of humans' instinct of committing intervention (punishment in particular) in response to severe transgression, as characterized by the preference for punitive over compensatory decisions in the face of criminal offenses (Van Prooijen, 2010). These findings coincide with Gintis et al.' (2003) concept of strong reciprocity, where individuals are inclined to punish norm violators or help others, even at a personal cost, without expecting any repayment. It is worth noting that the rate of punishment decision, but not that of compensation, increased during highly compared to moderately unfair events. A highly unfair allocation represents a severe transgression of the norm of justice. Hence, increased punishment may exert a greater impact in restoring equity, an effect that may be absent for compensation, as postulated by Van Doorn et al. (2018).

4.2. Pre-intervention: the mediating role of negative emotions

Based on the results that event unfairness significantly influenced intervention behavior, we analyzed the role of emotional states during this influence, and observed that negative emotion triggered by unfair events mediated the influence of event unfairness on intervention behavior. Meanwhile, participants who chose to intervene experienced more negative emotions compared to those who opted for inaction. Unlike previous studies that evoked incidental emotions unrelated to unfair events (e.g. incidental anger/happiness, Vayness et al., 2020; Gummerum et al., 2016), the present study measured participants' natural emotional experience upon perception of unfair events during a TPI task, which allows participants to freely decide whether and how to intervene. Indeed, the results revealed that inherent negative emotions induced by unfair events serve as a direct trigger of intervention selection and punishment decision, as shown by the significant mediation effects of emotion (see Fig. 3). This result is supported by the findings of Pfattheicher et al. (2019) that a prosocial emotion of compassion for the suffering victims increased third-party punishment. Future research should further explore whether task-irrelevant incidental emotions simply amplify the effect of inherent emotion triggered by unfair events, or they influence third-party intervention behavior in a different route.

Moreover, the results highlighted the crucial role of emotion in driving different intervention methods, as the model revealed a mediation effect of emotional state in the association between event unfairness and punishment selection. In our study, individuals exhibited stronger negative emotions during more unfair events, and their preferences for intervention methods shifted from compensation to punishment as unfairness increased from moderate to high. This supports Van Doorn and Brouwers's (2020) perspective that increased unfairness intensifies negative emotions in third parties, which ultimately shifts

their intervention preference toward punishment. This effect may stem from the severely unfair event that leads to attentional focus on the perpetrator (David et al., 2017), which contributed to a preference for offender punishment over victim compensation during severe offense (Van Prooijen, 2010). It should be noted that a couple of studies reported a compensation preference over punishment in TPI task (Dhaliwal et al., 2021; Van Doorn & Brouwers, 2020), despite greater weight of punishment in justice restoration being assumed (Van Doorn et al., 2018). Though these studies used various fairness-violating scenarios, they did not consider the manipulation of the severity of unfairness. Actually, we indeed observed a compensation preference during moderately unfair events, consistent with these prior studies. As such, the current study provides important insight into future studies of TPI, that the severity of unfair events subjectively perceived by individuals should be taken into consideration for its potential effect in shifting our intervention preference.

However, the present study used an integral emotion rating scale, which focuses on the valence and intensity of emotional experience without differentiating discrete emotion categories from one another. Prior studies have reported that experiencing the discrete emotion of anger promotes justice-maintaining action or third-party punitive behavior (Gummerum et al., 2016; Van de Vyver & Abrams, 2015). In this regard, whether the current observation, namely, negative emotion triggered by unfair events mediates their effects on third-party intervention/punishment, is a phenomenon applicable to all negative emotions or specific to injustice-induced anger remains an open question, which should be addressed in future studies.

4.3. Post-intervention: the emotional benefit of the intervention

We further examined emotional alterations following the intervention. Although TPI is typically viewed as not bringing material benefits, our findings indicate that participants experienced emotional improvement through their intervention behaviors. This aligns with the view that acts of kindness can enhance the emotional well-being of the actor (Aknin et al., 2020; Curry et al., 2018; Hui et al., 2020). Our results extend these prior studies by showing that this effect, namely, prosocial action for emotional benefit, also applies in unfairness-triggered TPI situations involving self-sacrifice. This inference is also supported by the current findings that participants who chose to intervene exhibited a more negative emotional state than those who chose inaction in pre-decision and also, the former received more emotional improvement post intervention.

We observed that the emotional improvement effect post versus pre intervention was influenced by intervention ability. While all groups experienced emotional improvement after intervention, participants with strong intervention ability showed greater emotional improvement than those in other groups. Their emotional state even shifted to a positive level, while those with weaker intervention abilities remained in a negative state. One possible explanation is that prosocial spending is most fulfilling when individuals can see the positive impact of their actions (Dunn et al., 2014). Participants with strong intervention abilities are likely to experience the effects of their behavior more directly than others. Previous research also highlights the strong link between power and emotion, suggesting that individuals with power often experience increased confidence, which enables them to feel more capable of regulating their negative emotions and managing them effectively (Zerwas et al., 2023). In our study, participants with strong intervention abilities were able to achieve their desired outcomes more easily, which enhanced their sense of power and self-efficacy, leading to a more positive emotional state. In contrast, participants with weaker intervention abilities experienced emotional relief from helping, which, however, was not as satisfactory as that in the former group. Additionally, participants experienced greater emotional improvement after intervention in highly unfair situations. This may be due to that participant perceived their intervention as helping others in times of need during highly unfair events, which amplified their emotional improvement.

Moreover, we found that compensation was more effective at improving emotional states than punishment. This is perhaps because compensation and punishment stem from different justice motives: compensation aligns with a restorative justice motive, aiming to soothe the victim and is often driven by prosocial emotions like sympathy (Engeler & Raihani, 2024). Whereas punishment reflects a retributive justice motive, focusing on the perpetrator, and is frequently fueled by anger (Van Prooijen, 2010). Given these distinct psychological pathways, the emotional superiority of compensation likely has boundary conditions.

First, external factors like the perceived severity of the transgression may influence the emotional improvement of the intervention. Our results showed that participants' preference shifted from compensation to punishment as unfairness intensified, suggesting that in the face of severe violations, the emotional calculus may also shift, potentially making punishment more emotionally gratifying. Weiß et al. (2023) argue that the preference for compensation is more stable in live interaction paradigms due to the 'observer effect', which may lead third parties to favor compensation as a means of signaling empathy and gaining social approval. Specifically, third parties are more inclined to compensate victims when their actions are publicly observable, but tend to punish offenders when interventions occur in private contexts. This may be because reputational outcomes play a crucial role in shaping third-party intervention preferences. Those who choose to compensate victims are evaluated more favorably than those who impose punishment (Li et al., 2021). Second, personal traits likely modulate the emotional benefits of these intervention methods. For instance, higher levels of trait empathy may enhance the affective rewards of compensatory actions, as such behaviors align more closely with an individual's moral self-concept. Conversely, trait anger may influence the emotional consequences of punitive behaviors. Individuals with higher trait anger might experience stronger emotional relief when engaging in punitive actions that restore a sense of fairness. Considering that no prior research has directly examined the emotional consequences of different intervention strategies, future studies should explicitly test these interactions to clarify how contextual factors and personal traits jointly shape the affective outcomes of third-party interventions.

Notably, third parties who chose inaction also experienced emotional improvement in the post-decision. Given that a lack of motivation can lead individuals to choose different methods for emotion regulation (Yang, Yan, et al., 2022), a feasible explanation is that those who opt for inaction may do so to protect their self-interest. As such, the option of inaction may have served as an effective strategy for emotional self-protection in contrast to the prosocial TPI decision.

4.4. Limitations

Several limitations need to be acknowledged. Firstly, as described earlier, the current study didn't differentiate discrete categories of emotion related to injustice perception, which made it unlikely to ascertain which specific emotion(s) mediate the effect of unfair events on TPI behavior. Thus, the role of specific emotion categories in the effect of unfair events on TPI behavior, both intervention/inaction decision and compensation/punishment selection, needs to be directly examined in further studies. Secondly, to date, the studies of the relations between emotion and TPI either induced incidental, task-irrelevant emotional state (Gummerum et al., 2016; Vayness et al., 2020) or manipulated emotion integral to unfairness-related TPI context (Inzlicht et al., 2015; Pfattheicher et al., 2019). Henceforth, how pre-existing, incidental emotional context interacts with integral emotion elicited by unfairness in shaping TPI is of applied significance, thus requires further investigation. Thirdly, the present study focused on the cognitive pathways and the consequences of third-party intervention during unfair events, from the perspective of injustice-related emotions.

In order to better explain the source of emotional improvement, a post-decision question regarding justice perception should have been included. This limitation should be overcome in forthcoming lines of studies. In line with this, neurobiological mechanisms accompanying the mediating role of emotion, and those subserving emotional improvement post intervention remain elusive, awaiting in-depth investigation in future studies.

4.5. Conclusions

By investigating the dynamic changes of emotion during TPI, the present study highlights the crucial role of emotional improvement in self-sacrificial altruistic behavior, offering a new avenue for understanding self-sacrifice altruistic behavior. The results show that negative emotional state triggered by unfair events mediates the effect of event unfairness on individuals' intervention behavior. More unfair events lead to more frequent and intense interventions, shifting individuals' preferences from compensation to punishment. Furthermore, emotional improvement occurs after intervention, and those possessing stronger intervention abilities or facing more unfair events experienced greater emotional improvement. Additionally, compensatory interventions lead to greater emotional improvement than punitive ones.

Clinical trial registration

No.

CRediT authorship contribution statement

Jiachen Qu: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jan Drewes:** Writing – review & editing, Supervision, Investigation. **Yunfeng He:** Writing – review & editing, Formal analysis. **Jiajin Yuan:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Ethics approval statement

All procedures in this study were conducted in accordance with the principles outlined in the Declaration of Helsinki 2004 and its subsequent amendments, or comparable ethical standards. The experiment was approved by the local Ethics Committee of Sichuan Normal University. All participants provided written informed consent prior to their participation in the experiment.

Permission to reproduce material from other sources

No.

Funding statement

This study was supported by the National Natural Science Foundation of China (NSFC31971018), the Human and Social Science Fund of Ministry of Education of China (24XJA190003), and the Distinguished Young Scholar Fund in Sichuan province (2023NSFSC1938).

Declaration of competing interest

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.actpsy.2025.106119.

Data availability

The data are available in Mendeley Data and are accessible at https://data.mendeley.com/datasets/zsr4ts2g5c/1.

References

Aknin, L. B., Dunn, E. W., Proulx, J., Lok, I., & Norton, M. I. (2020). Does spending money on others promote happiness?: A registered replication report. *Journal of Personality and Social Psychology, 119*(2), Article e15. https://doi.org/10.1037/pspa0000191

Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of Personality and Social Psychology, 107*(3), 393. https://doi.org/10.1037/a0037207

Berdahl, J. L., & Martorana, P. (2006). Effects of power on emotion and expression during a controversial group discussion. *European Journal of Social Psychology, 36*(4), 497–509. https://doi.org/10.1002/ejsp.354

Bergquist, M., & Ekelund, M. (2025). The role of emotion regulation in normative influence under uncertainty. *BMC Psychology, 13*(1), 1–15. https://doi.org/10.1186/s40359-025-03033-z

Berman, J. Z., Levine, E. E., Barasch, A., & Small, D. A. (2015). The Braggart's dilemma: On the social rewards and penalties of advertising prosocial behavior. *Journal of Marketing Research, 52*(1), 90–104. https://doi.org/10.1509/jmr.14.0002

Buttrick, N. R., Heintzelman, S. J., & Oishi, S. (2017). Inequality and well-being. *Current Opinion in Psychology, 18*, 15–20. https://doi.org/10.1016/j.copsyc.2017.07.016

Buttrick, N. R., & Oishi, S. (2017). The psychological consequences of income inequality. *Social and Personality Psychology Compass, 11*(3), Article e12304. https://doi.org/10.1111/spc3.12304

Cheng, X., Zheng, L., Liu, Z., Ling, X., Wang, X., Ouyang, H., … Guo, X. (2022). Punishment cost affects third-parties' behavioral and neural responses to unfairness. *International Journal of Psychophysiology, 177*, 27–33. https://doi.org/10.1016/j.ijpsycho.2022.04.003

Curry, O. S., Rowland, L. A., Van Lissa, C. J., Zlotowitz, S., McAlaney, J., & Whitehouse, H. (2018). Happy to help? A systematic review and meta-analysis of the effects of performing acts of kindness on the well-being of the actor. *Journal of Experimental Social Psychology, 76*, 320–329. https://doi.org/10.1016/j.jesp.2018.02.014

David, B., Hu, Y., Krüger, F., & Weber, B. (2017). Other-regarding attention focus modulates third-party altruistic choice: An fMRI study. *Scientific Reports, 7*(1), Article 43024. https://doi.org/10.1038/srep43024

Davis, M. H. (1980). *A multidimensional approach to individual differences in empathy*.

Dhaliwal, N. A., Patil, I., & Cushman, F. (2021). Reputational and cooperative benefits of third-party compensation. *Organizational Behavior and Human Decision Processes, 164*, 27–51. https://doi.org/10.1016/j.obhdp.2021.01.003

Dunn, E. W., Aknin, L. B., & Norton, M. I. (2014). Prosocial spending and happiness: Using money to benefit others pays off. *Current Directions in Psychological Science, 23*(1), 41–47. https://doi.org/10.1177/0963721413512503

Engeler, N. C., & Raihani, N. J. (2024). Acting without considering personal costs signals trustworthiness in helpers but not punishers. *Communications Psychology, 2*(1), 47. https://doi.org/10.1038/s44271-024-00092-7

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149

Fehr, E., & Fischbacher, U. (2004a). 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., & Fischbacher, U. (2004b). Third-party punishment and social norms. *Evolution and Human Behavior, 25*(2), 63–87. https://doi.org/10.1016/s1090-5138(04)00005-4

FeldmanHall, O., Dalgleish, T., Evans, D., & Mobbs, D. (2015). Empathic concern drives costly altruism. *Neuroimage, 105*, 347–356. https://doi.org/10.1016/j.neuroimage.2014.10.043

Ginther, M. R., Hartsough, L. E., & Marois, R. (2022). Moral outrage drives the interaction of harm and culpable intent in third-party punishment decisions. *Emotion, 22*(4), 795. https://doi.org/10.1037/emo0000950

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24*(3), 153–172. https://doi.org/10.1016/s1090-5138(02)00157-5

Guinote, A. (2017). How power affects people: Activating, wanting, and goal seeking. *Annual Review of Psychology, 68*, 353–381. https://doi.org/10.1146/annurev-psych-010416-044153

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. https://doi.org/10.1016/j.jesp.2016.04.004

Hu, Y., Strang, S., & Weber, B. (2015). Helping or punishing strangers: Neural correlates of altruistic decisions as third-party and of its relation to empathic concern. *Frontiers in Behavioral Neuroscience, 9*, 24. https://doi.org/10.3389/fnbeh.2015.00024

Huang, L., Yang, T., & Li, Z. (2003). Applicability of the positive and negative affect scale in Chinese. *Chinese Mental Health Journal*. https://doi.org/10.3321/j.issn:1000-6729.2003.01.018

Hui, B. P., Ng, J. C., Berzaghi, E., Cunningham-Amos, L. A., & Kogan, A. (2020). Rewards of kindness? A meta-analysis of the link between prosociality and well-being. *Psychological Bulletin, 146*(12), 1084. https://doi.org/10.1037/bul0000298

Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences, 19*(3), 126–132. https://doi.org/10.1016/j.tics.2015.01.004

Jordan, J. J., Hoffman, M., Bloom, P., & Rand, D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature, 530*(7591), 473–476. https://doi.org/10.1038/nature16981

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. https://doi.org/10.1177/0956797615624466

Landmann, H., & Hess, U. (2017). What elicits third-party anger? The effects of moral violation and others' outcome on anger and compassion. *Cognition and Emotion, 31*(6), 1097–1111. https://doi.org/10.1080/02699931.2016.1194258

Leliveld, M. C., van Dijk, E., & van Beest, I. (2012). Punishing and compensating others at your own expense: The role of empathic concern on reactions to distributive injustice. *European Journal of Social Psychology, 42*(2), 135–140. https://doi.org/10.1002/ejsp.872

Li, Z., Hu, G., Xu, L., & Li, Q. (2021). Third-party punishment or compensation? It depends on the reputational benefits. *Frontiers in Psychology, 12*, Article 676064. https://doi.org/10.3389/fpsyg.2021.676064

O'Reilly, J., Aquino, K., & Skarlicki, D. (2016). The lives of others: Third parties' responses to others' injustice. *Journal of Applied Psychology, 101*(2), 171–189. https://doi.org/10.1037/apl0000040

Pfattheicher, S., Sassenrath, C., & Keller, J. (2019). Compassion magnifies third-party punishment. *Journal of Personality and Social Psychology, 117*(1), 124. https://doi.org/10.1037/pspi0000165

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods, 16*(2), 93. https://doi.org/10.1037/a0022658

Qu, C., Tang, Z., Zhang, H., Hu, Y., & Dreher, J. C. (2018). Neural modulation of social distance on third-party punishment. *BioRxiv*. , Article 274720. https://doi.org/10.1101/274720

Quoidbach, J., Mikolajczak, M., & Gross, J. J. (2015). Positive interventions: An emotion regulation perspective. *Psychological Bulletin, 141*(3), 655. https://doi.org/10.1037/a0038648

Van de Vyver, J., & Abrams, D. (2015). Testing the prosocial effectiveness of the prototypical moral emotions: Elevation increases benevolent behaviors and outrage increases justice behaviors. *Journal of Experimental Social Psychology, 58*, 23–33. https://doi.org/10.1016/j.jesp.2014.12.005

Van Doorn, J., & Brouwers, L. (2020). Third-party responses to injustice: A review on the preference for compensation. *Crime Psychology Review, 3*(1), 59–77. https://doi.org/10.4324/9780429346927-18

Van Doorn, J., Zeelenberg, M., & Breugelmans, S. M. (2018). An exploration of third parties' preference for compensation over punishment: Six experimental demonstrations. *Theory and Decision, 85*, 333–351. https://doi.org/10.1007/s11238-018-9665-9

Van Prooijen, J. W. (2010). Retributive versus compensatory justice: Observers' preference for punishing in response to criminal offenses. *European Journal of Social Psychology, 40*(1), 72–85. https://doi.org/10.1002/ejsp.611

Vayness, J., Duong, F., & DeSteno, D. (2020). Gratitude increases third-party punishment. *Cognition and Emotion, 34*(5), 1020–1027. https://doi.org/10.1080/02699931.2019.1700100

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063. https://doi.org/10.1037/0022-3514.54.6.1063

Weiß, M., Saulin, A., Iotzov, V., Hewig, J., & Hein, G. (2023). Can monetary incentives overturn fairness-based decisions? *Royal Society Open Science, 10*(6), Article 211983. https://doi.org/10.1098/rsos.211983

Yang, J., Gu, R., Liu, J., Deng, K., Huang, X., Luo, Y. J., & Cui, F. (2022). To blame or not? Modulating third-party punishment with the framing effect. *Neuroscience Bulletin, 38*(5), 533–547. https://doi.org/10.1007/s12264-021-00808-3

Yang, J., Yan, X., Chen, S., Liu, W., Zhang, X., & Yuan, J. (2022). Increased motivational intensity leads to preference for distraction over reappraisal during emotion regulation: Mediated by attentional breadth. *Emotion, 22*(7), 1595. https://doi.org/10.1037/emo0000977

Zerwas, F. K., Tharp, J. A., Chen, S., & Mauss, I. B. (2023). Individual differences in social power: Links with beliefs about emotion and emotion regulation. *Journal of Personality, 91*(2), 314–331. https://doi.org/10.1111/jopy.12721

Zhang, F. F., Dong, Y., & Wang, K. (2010). Reliability and validity of the Chinese version of the Interpersonal Reactivity Index-C. *Chinese Journal of Clinical Psychology*. https://doi.org/10.3969/j.issn.1005-9202.2014.20.130