Social Cognitive and Affective Neuroscience, 2025, **20**(1), nsaf042
DOI: https://doi.org/10.1093/scan/nsaf042
Advance Access Publication Date: 2 May 2025
Original Research – Neuroscience

Altruistic or fair? The influence of empathy on third-party punishment: an event-related potential study

Guanfei Zhang[1,2], Min Tan[1,2], Jin Li[1,2,*], Yiping Zhong[1,2,*]

[1]Department of Psychology, Hunan Normal University, Changsha, Hunan 410081, China
[2]Cognition and Human Behavior Key Laboratory of Hunan Province, Hunan Normal University, Changsha, Hunan 410081, China
[*]Corresponding author. Department of Psychology, Hunan Normal University, No. 36 Lushan Road, Yuelu Dist, Changsha, Hunan, 410081, China.
E-mail: jin.li@hunnu.edu.cn; Yiping Zhong, E-mail: ypzhong@hunnu.edu.cn

Abstract

Although most individuals strongly prefer social fairness and punish behaviours that violate fairness norms, recent psychological studies have shown that empathy towards 'perpetrators' who violate fairness norms can affect people's fairness decision-making, resulting in tolerance for unfair behaviour, even as direct 'victims' of unfair behaviour. However, in real life, people more often view unfair events from a third-party perspective, and little is known about how empathy affects fairness decisions by third parties whose self-interests are not threatened and their neurocognitive mechanisms. The present study examined effects of empathy directed towards a 'perpetrator' on third-party punishment using event-related potentials. The results suggest that, in the nonempathy condition, unfair offers induced stronger unfairness aversion in third-party decision makers and increased motivation and cognitive resource investment to alleviate this negative emotion compared to fair offers, reflecting the greater amplitude differences of fairness effects on the anterior N1 component, medial frontal negative, and smaller late positive components in the nonempathy condition. However, in the empathy condition, the differential impact of the fairness effect disappeared. These findings reveal the neural basis for trade-offs between altruistic and fairness motives in third-party fairness decision-making processes involving empathy.

Keywords: empathy; third-party punishment; event-related potential; anterior negativity-1; medial frontal negativity; late positive component

Introduction

Fairness is an important social norm that can promote group prosperity and social stability; punishing deviations from this norm is an effective way to maintain long-term social order and promote interpersonal cooperation and communication (Luo et al. 2024). For example, regulatory agencies and judges restore social justice by providing proportional punishments to perpetrators based on the seriousness of their offenses (Kundro et al. 2023). However, sometimes the perpetrator may elicit empathy from the public by detailing personal harms they have experienced in an attempt to weaken just punishment for themselves (Haegerich and Bottoms 2000, Tsoudis 2002). Researchers and practitioners in the field of justice also acknowledge that empathy inevitably affects judges' and juries' decisions (Hoffman 2008, Gu et al. 2025).

According to the empathy–altruism hypothesis, empathy is associated with concern for the interests of others (Decety 2015, Decety et al. 2016). Awakening an individual's empathy activates altruistic motivation, which in turn leads to appropriate altruistic decisions and behaviours (Batson et al. 1991, 2015, Batson and Shaw 1991). Therefore, the altruistic motive arising from empathy towards the perpetrator may affect the fairness motive of the third party imposing punishment, thus reducing the degree of punishment imposed on the perpetrator. For example, perpetrators of harm commonly use news media reports to claim that the act of harming others was caused by trauma from family and/or school bullying in childhood, to induce empathy and altruistic motives in the public in an attempt to interfere with judicial fairness (Ding and Javed 2020). Considering the popularity of the Internet and 'clickbait channels', the public's media network participation is extremely high. If the public's understanding of this phenomenon is not sufficiently clear, it is likely to be intentionally misused, resulting in conformity and blind following, interfering with the fairness of judicial decisions with public opinion, and even calling the credibility of the judiciary into question (Picó et al. 2020). Therefore, the mechanism of

Received: 29 October 2024; **Revised**: 28 March 2025; **Accepted**: 29 April 2025

this phenomenon needs to be explored, which will help the public view violations of social norms from a rational and peaceful third-party perspective and assist the government in formulating intervention strategies accordingly. However, the impact of altruistic motives arising from empathy towards perpetrators on third-party punitive behaviours involving fairness motives and the underlying neural mechanisms behind their effects are still unknown.

Previous studies have shown that altruistic motivation generated by empathy can affect fairness in decision-making processes. For example, inducing supervisors to empathize with group members affects the fairness of their task assignments, as evidenced by supervisors assigning positive tasks to members experiencing difficulties (empathy condition) and negative tasks to members not experiencing difficulties (nonempathy condition), rather than using a fair practice of flipping a coin (Batson et al. 1995). In the Ultimatum Game (UG), tasks involving a second-party perspective are used to measure an individual's acceptance or rejection of a fair or unfair offer of allocation for a known amount (Güth, et al., 1982, Jaroslawska et al. 2020, He et al. 2022, Wang et al. 2022). It was found that when the participants were the recipients of the UG, that is, when they played the role of 'victims' of unfair behaviour, empathy with the allocator (e.g. left-behind students) increases their tolerance for unfair offers by the allocator (He et al. 2022, Wang et al. 2022). Researchers explain this as individuals prioritizing the use of empathy-altruistic motives to guide behaviour when faced with conflicts between empathy-altruistic and fairness motives (He et al. 2022). This also means that in fair decision-making, if an empathy-altruistic motivation towards the 'perpetrator' is induced, altruistic motivation tends to surpass fairness motivation in the conflict balance between the two motivations (Wang et al. 2022), and empathy-altruistic motivation interferes with fairness decision-making. Even if the participants are the 'victims' of unfair behaviour, their empathy-altruistic motives towards the 'perpetrator' still have a certain negative impact on fairness decision-making (He et al. 2022, Wang et al. 2022). However, in these studies, these individuals themselves were either 'perpetrators' of unfair behaviour due to empathy-altruistic motives, or 'victims' being treated unfairly, while in real life, people tend to view a perpetrator's unfair behaviour towards a victim from a third-party perspective rather than being directly involved. It remains unknown whether third-parties whose own interests are not threatened empathize with perpetrators of unfair behaviour, similar to the 'victims' as second parties, and exhibit a similar tendency to prioritize empathy-altruistic motives over fairness motives.

A paradigm used to explore fairness decision-making involving third parties whose own interests are not threatened is the third-party punishment (TPP) paradigm. TPP refers to the tendency of individuals, as disinterested third-party bystanders who have not been directly affected by illegal behaviour, to punish violators in order to uphold social norms when they observe such behaviour (Cui et al. 2019, Gao et al. 2025). Simultaneously, TPP behaviour requires individuals to pay for their own costs without receiving significant rewards (Jordan et al. 2016, Zheng et al. 2024). The TPP paradigm in the laboratory was adapted from the dictator game (DG) (Cui et al. 2019). In the classic DG, the allocator has the right to randomly allocate a sum of money, while the receiver can only accept the amount allocated by the allocator, with no right to refuse, regardless of fairness (Gummerum et al. 2010). Allocating half of the amount to the recipient by the allocator is considered a fair offer, whereas distributing less than half of the total amount is considered an unfair offer or a violation of social norms (Sun et al. 2015, Cui et al. 2019). The TPP paradigm adds to the DG a neutral third party with the right to punish allocators for unfair behaviour by incurring its own monetary costs (Cui et al. 2019, Ouyang et al. 2021, Zheng et al. 2024), a paradigm that can effectively reflect the third party's attitude towards fairness norms (Fehr and Fischbacher 2004). In the TPP paradigm involving fairness norms, when an allocator makes an unfair allocation, third-party individuals experience inequality aversion and negative emotions towards the violation of fairness norms (Fehr and Gächter 2002, Li et al. 2023b, Treadway et al. 2014, Yang et al. 2022, Zheng et al. 2024). In turn, third parties voluntarily give up their self-interest to punish unfair behaviour to pursue fairer outcomes (Sun et al. 2015, Zinchenko et al. 2023). Therefore, the motivation for TPP decisions is not self-interest but rather driven by inequality aversion and negative emotions, to maintain fairness norms (Yin et al. 2019, Zheng et al. 2024).

Importantly, previous studies have focused on the role of empathy in TPP decisions, which involve the absence of threats to one's own interests. For example, research has found a positive correlation between trait empathy in third-party observers and feedback-related negativity (FRN) amplitude in TPP decisions (Mothes et al. 2016). That is, the lower the trait empathy ability of third-party observers, the more negative the FRN amplitude they exhibited in unfair allocation trials. FRN components are associated with anticipatory violations (Holroyd and Krigolson 2007). Researchers speculate that this may be because individuals with higher empathy abilities are better able to view unfair distributions from the perspective of allocators and, therefore, this result is not surprising (Mothes et al. 2016). These studies focused on the relationship between individual differences in empathy traits of third-party individuals and TPP, but it is not clear how empathy for the 'perpetrator' will affect TPP decisions. On the one hand, these studies focused on the correlation between empathy traits in third-party observers themselves and TPP, rather than manipulating empathy towards a 'perpetrator' and conducting causal research on TPP. On the other hand, the empathy trait in these studies was that of third-party individuals, without empathy oriented directionality, i.e. it is not empathy towards the 'perpetrator' who make unfair allocations. Therefore, it is necessary to refine the empathy orientation and explore its impact on TPP. Although some studies have controlled the directionality of empathy, they manipulated the empathy of third-party observers towards victims, and found that the stronger their empathy towards victims, the stronger the punishment for the 'perpetrator' who caused this result (Pfattheicher et al. 2019). However, when an individual's empathy is directed differently, such as towards the perpetrator or victim, their behavioural decisions and brain processing mechanisms will exhibit different characteristics (Hernandez Pena et al. 2024). Therefore, this study manipulated the empathy orientation of third-party individuals, focusing on empathy towards a 'perpetrator' (i.e. unfair allocators in TPP), and explored the role of this empathy orientation in fair decision-making using the TPP paradigm.

More importantly, to better understand the role of individual empathy for perpetrators in TPP, this study used event-related potentials (ERPs) to reveal individual decision-making processes within a time-locked period. According to previous ERP research, the time processing of fair decision-making involves three stages: early automated intuitive processing, middle stage semi-automatic emotional processing, and late-stage cautious cognitive processing (Wang et al. 2014, Hu and Mai 2021, He et al. 2022, Chen et al. 2024). First, the ERP component associated with early automated intuitive processing in fair decision-making is the anterior N1 component (AN1) (Wang et al. 2014, He et al. 2022, Chen et al. 2024). AN1 is a negative wave that peaks in the frontal regions ∼100 ms after the presentation of a visual stimulus, indicating early attention processing towards the stimulus (Luck 1995). AN1 is associated with recognition of unexpected stimuli in social decision processing (Annic et al. 2014, Duzcu et al. 2019, Bouwer et al. 2020) or the violation of anticipated activities (He et al. 2022). Since people typically seek fairness in society and adhere to principles of equity (Peters and van den Bos 2008, Hu et al. 2022), unfair offers are less in line with individual expectations and induce larger AN1 amplitudes than fair offers (Wang et al. 2014, He et al. 2022). Second, the ERP component related to semi-automated emotional processing in fairness decision-making was medial frontal negativity (MFN). MFN is a negative wave that originates in the anterior cingulate cortex and reaches its peak 250–350 ms after stimulus onset (Gehring and Willoughby 2002, Wang et al. 2014). MFN is similarly sensitive to violations of social norms, and unfair offers that violate fairness norms generate larger MFN amplitudes than fair offers (Wang et al. 2016, Zinchenko et al. 2023). Finally, the ERP component related to late-stage prudent cognitive processing in fairness decision-making was the late positive component (LPC) (Cui et al. 2019, Ouyang et al. 2021, Zheng et al. 2024, Zheng et al. 2024;). The LPC is a positive wave that peaks ∼400–600 ms after stimulus onset, reflecting participation in executive control during the decision-making process; LPC amplitude decreases with increasing demands for executive control (Johnson et al. 2003, 2008). In TPP decisions, facing an unfair offer requires more cognitive resources and results in a smaller LPC amplitude than facing a fair offer (Cui et al. 2019, Ouyang et al. 2021). Taken together, these findings indicate that the AN1, MFN, and LPC are sensitive to allocation fairness.

Therefore, the current study utilized the TPP paradigm and manipulated empathy towards unfair allocators to explore the impact of empathy on TPP. Behaviourally, according to previous studies (Batson et al. 1995, He et al. 2022), empathy-altruistic motivation can negatively affect motivation to behave fairly. For example, when participants are the direct victims of unfair allocation, empathy towards the ‘perpetrator’ of unfair allocation increases individuals’ tolerance of unfair allocation. Specifically, in cases of unfair allocation, participants with empathy have higher acceptance rates than those without empathy (He et al. 2022). Therefore, we predict that in TPP, when faced with unfair allocation, participants, as third-party observers, who empathize with the allocator will have a lower punishment rate and amount than those who do not. Meanwhile, neurally, we mainly focused on the three ERP components of fair decision processing. Therefore, we assume that unfair allocation will result in larger AN1 and MFN amplitudes and smaller LPC amplitudes than fair allocation in nonempathy conditions.

Methods

Participants and experimental design

A power analysis (G*Power 3.1, power = 0.80, effect size $f = 0.25$, $\alpha = 0.05$) suggested that 24 participants would ensure 80% statistical power in the case of small to medium effect sizes (Faul et al. 2007). We recruited 33 healthy undergraduate or graduate students, 2 of whom were excluded from the analysis because their behavioural data exceeded ±3 standard deviations (SD). Thus, 31 participants (16 female and 15 male, mean age = 20.48, $SD = 2.00$, range = 18–25 years) from Hunan Normal University participated in this experiment, which had a 2 (empathy: empathy vs. nonempathy) × 2 (fairness: fair vs. unfair) within-participants factorial design (He et al. 2022, Wang et al. 2022). All participants were right-handed, had normal or corrected-to-normal vision, and signed informed consent before the experiment. A certain reward would be given after the experiment. This study was approved by the Ethics Committee of the Department of Psychology, Hunan Normal University (No. 2023-287).

Experimental materials

Empathy manipulation

Empathy manipulation refers to empathy directed towards the ‘perpetrator’; therefore in the TPP paradigm, empathy is directed towards the proposer. Referring to previous research, we manipulated the empathy condition as the proposer being left-behind students from a remote rural school, while the nonempathy condition was manipulated as the proposer being ordinary students from a normal urban school (Liu et al. 2018, 2020, He et al. 2022, Wang et al. 2022). Compared to ordinary students, left-behind students have relatively poor family conditions and school infrastructure. Present background information was required participants to read thoroughly to induce individual empathy for left-behind students (Liu et al. 2018, 2020, He et al. 2022, Wang et al. 2022). The effectiveness of empathy materials was evaluated through preliminary experiments, and they could effectively induce an individual’s empathy state, see supplementary for details.

Fairness perception assessment and emotional assessment

Negative emotions induced by inequality aversion and unfair distribution can affect individuals’ decision-making in TPP (Cui et al. 2019, Zinchenko et al. 2023, Li et al. 2023b). Therefore, we also measured individuals’ perception of fairness and related emotions to understand the mechanisms by which individuals make TPP decisions with or without empathy (He et al. 2022). The participants were asked to assess emotional feelings with a 9-point scale (5 = Calm mood, ‘6–9’ = ‘The larger the number, the more positive the emotion’, ‘1–4’ = ‘The smaller the number, the more negative the emotion’) (Bradley et al. 2001, Hewig et al. 2011, Hu and Mai 2021, He et al. 2022) and assess perceived fairness with 7-point scale (1 = very unfair, 4 = uncertain, 7 = Very fair) (Pillutla and Murnighan 1996, Leliveld et al. 2012, He et al. 2022). Throughout the entire experiment, there were four blocks under both empathy and nonempathy conditions, and evaluations of the four allocation schemes, fair offers (5:5 and 6:4) and unfair offers (8:2 and 9:1), were randomly generated in each block. Therefore, participants were required to undergo four evaluations of fairness perception evaluation and emotion evaluation under each block under empathy and nonempathy conditions.

TPP task

The TPP paradigm was based on the UG paradigm and incorporated a third party. The participants acted as a third party to observe the DG between the allocator (first party) and the recipient (second party) (Cui et al. 2019, Ouyang et al. 2021). In the DG, there were 10 tokens in each round, and the allocator could freely allocate 10 tokens between themselves and the recipient, and the recipient couldn’t refuse regardless of the allocation result. The participant watched the game from a third-party perspective and received five tokens in each round of the game. These five tokens could be kept for themselves or used to punish the allocator. When punished, for each token taken out by the participant, the allocator could be reduced by two tokens (Cui et al. 2019). The punishment options included five options from 1 to 5, which could reduce the allocator by 2/4/6/8/10 tokens respectively. 1 token represented 1 RMB in the final participant fee settlement. The settlement of the participant fee was 20 RMB basic fee + (the remaining token cost randomly selected from 10 rounds of results * 1 RMB), which meant the fee ranged from 20 to 70 RMB (Cui et al. 2019).

In order to make the experiment more credible, the following content would be emphasized to the participants. First, in order to reduce the negative emotions generated by participants when making unfair distributions to a single allocator and their subsequent punishment decisions, the researchers would explain to the participants that the allocators in each round of the game were proposed by different people, as it was difficult to gather many people in a unified location for gaming. Therefore, these allocation proposals were collected in advance (He et al. 2022). Second, the researcher would emphasize to the participants that the punishment decisions they made would truly affect the benefits of the allocator. Because the allocators who collected the proposals in advance did not receive the trial fees, they needed to distribute the trial fees uniformly based on the participants' decisions (Ouyang et al. 2021). Finally, the allocator was unaware of the existence of a punitive agent when making their own proposal, and neither the allocator nor the recipient was aware of the punitive decision made by the participant (Cui et al. 2019, Ouyang et al. 2021).

Empathy rating

A simple 7-point scale was provided to rate participants' empathy for the two kinds of conditions (1 = I do not agree at all, 7 = I strongly agree) and participants were asked to describe to what extent they felt sympathetic/caring/touched (Pfattheicher et al. 2019, He et al. 2022). An internal consistency test was performed on the three measurement items for condition (empathy vs. nonempathy). Results showed that Cronbach's α coefficient was 0.731 (empathy condition) and 0.822 (nonempathy condition).

Experimental procedure

As illustrated in Fig. 1., each trial began with a fixation of 800–1000 ms random intervals followed by the information of allocators and recipients for 2000 ms. After a blank screen of 1000 ms, the allocator's offer was presented for 1500 ms. On the left is the allocator's own token, and on the right is the token given to the recipient by the allocator. Then a blank screen was presented for 1000 ms. Thereafter, decision interface was presented for 3000 ms, during which participants were required to decide on whether to keep or punish as rapidly as possible. If they wanted to keep, press the 'F' key. If they wanted to punish, press the 'J' key. And then a '1–5' punishment amount interface would appear, requiring participants to press the number button that they want to impose punishment on the allocator as soon as possible. After participants making the key decision, they would present an interval of 800–1000 ms, and then started the next trial.

During the experiment, 360 trials were divided into eight blocks (see Table 1); among these were fair offers (72 trials each for 6:4 and 5:5), unfair offers (72 trials each for 8:2 and 9:1), and filler offers (36 trials each for 7:3; not included in the data analysis) (He et al. 2022, Wang et al. 2022). The former represented the token received by the allocator, while the latter represented the token distributed by the allocator to the recipient. The experiment was divided into two tasks: empathy and nonempathy condition (He et al. 2022), with each task divided into four blocks and 45 trials in each block. The order of the experimental tasks was balanced. At each level of each block, emotional and fairness perception assessment were presented randomly (He et al. 2022). After the experiment, participants were required to complete empathy rating of different empathy proposers to further evaluate the effectiveness of empathy state manipulation in the experiment.

Electroencephalogram recording and analysis

Continuous Electroencephalogram (EEG) was recorded (band-pass: 0.01–100 Hz, sampling rate: 500 Hz) with ANT Neuro system, using a 64-lead electrode cap extended by the international 10–20 system. During EEG recording, all electrodes were online referenced to CPz first and then offline algebraic re-referenced to the average of left and right mastoids. The resistance of all electrodes remained <5 kΩ during the experiment. The EEG data were preprocessed and analysed using Matlab 2022b and EEGLAB toolbox (v.13.4.4 b, Swartz Center for Computational Neuroscience, La Jolla, CA, USA) according to certain methods. EEG

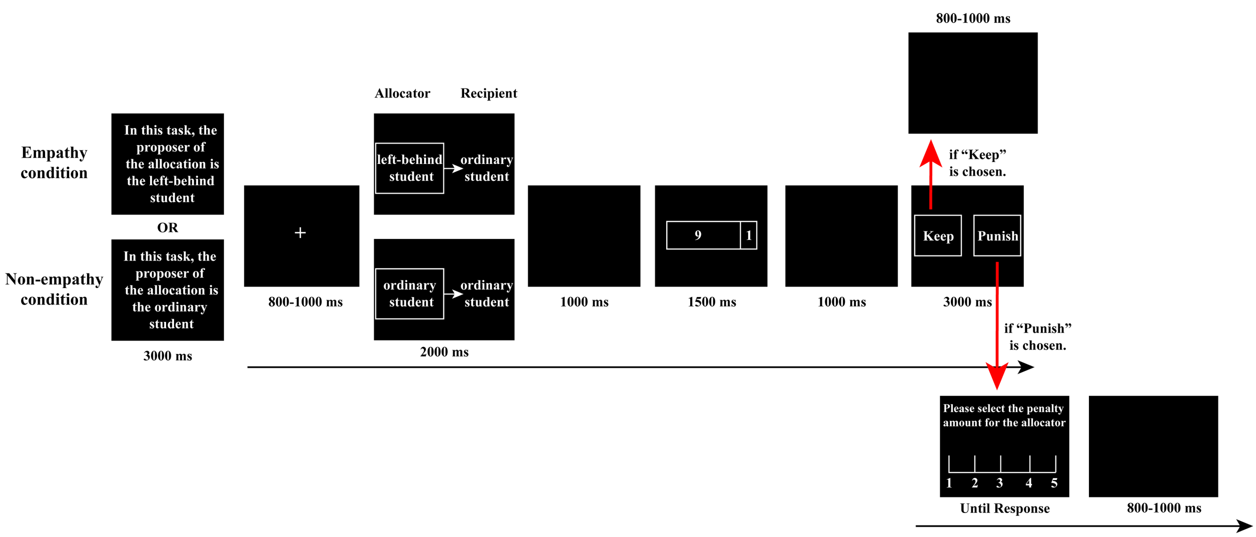

Figure 1. Illustration of a single trial of the experimental procedure.

Table 1. The number of trials under each condition.

Condition		Allocator: Recipient	Trials
Empathy	Unfair offers (72 trials)	9:1	36
		8:2	36
	Filler offers (36 trials)	7:3	36
	Fair offers (72 trials)	6:4	36
		5:5	36
Nonempathy	Unfair offers (72 trials)	9:1	36
		8:2	36
	Filler offers (36 trials)	7:3	36
	Fair offers (72 trials)	6:4	36
		5:5	36

data were filtered at 0.1–30 Hz (24 Db/oct). Independent component analysis was used to remove electro-oculogram blink and motion artifacts from EEG signals (Delorme and Makeig 2004, Plöchl et al. 2012). Then the EEG data were segmented into epochs from 200 ms before to 800 ms after the presentation of the offer interface. After baseline correction, epochs with more than ±100 μV artifacts were excluded from further analysis. Then, epochs were averaged separately for each condition of each participant.

We analysed average amplitudes of the AN1 (110–160 ms), MFN (290–330 ms), and LPC (380–480 ms) based on the grand-averaged ERP waveforms and prior studies (Cui et al. 2019, He et al. 2022, Wang et al. 2022). According to the topographical distribution (Figs 4b and 5b) and previous studies, we selected different sets of electrodes for each component. Specifically, the frontal AN1 and MFN were calculated as the average amplitude at F3, Fz, F4, FC3, FCz, and FC4; and LPC amplitudes was calculated as the average amplitude at F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, and CP4 (Cui et al. 2019, Zhang et al. 2020, He et al. 2022, Wang et al. 2022). All data were statistically analysed using SPSS 27.0. Each component and behaviour data were analysed using a 2 (empathy: empathy vs. nonempathy) × 2 (fairness: fair vs. unfair) repeated-measures analysis of variance (ANOVA) (Hu and Mai 2021, He et al. 2022, Wang et al. 2022, Chen et al. 2024). The significance level for all analyses was set at 0.05. *Post hoc* comparisons were Bonferroni-corrected at $P < 0.05$. The Greenhouse–Geisser correction was conducted to account for sphericity violations whenever appropriate, and the partial eta-squared (η_p^2) was reported as a measure of effect size.

Result

Behavioural results

Empathy manipulation checks

We performed a paired sample *t*-test for the empathy rating under the empathy and nonempathy conditions. Results found that the rating score under the empathy condition (mean ± standard: 6.16 ± 0.61) was significantly higher than the nonempathy condition (2.61 ± 1.00), t (30) = 16.33, $P < .001$, Cohen's $d = 2.93$, 95% CI [2.11, 3.75]. These results suggested that participants are more likely to empathize with proposers in the empathy condition than with proposers in the nonempathy condition. Hence, the manipulation of empathy was effective.

Punishment rate

Regarding the rate of punishment, the main effect of fairness was significant, F (1, 30) = 658.69, $P < .001$, $\eta_p^2 = 0.96$, indicating that the participants made more punishment decisions when facing unfair (0.91 ± 0.16) than fair offers (0.10 ± 0.11). The main effect of empathy was also significant, F (1, 30) = 27.54, $P < .001$, $\eta_p^2 = 0.48$, indicating that the participants made fewer punishment decisions in the empathy condition (0.42 ± 0.17) than in the nonempathy condition (0.59 ± 0.11). The interaction between fairness and empathy was not significant, F (1, 30) = 0.10, $P = .749$ (see Fig. 2).

Punishment amount

When the participants selected the 'keep' option, the punishment amount for that trial will be set to 0, which means the punishment amount was 0 (Liu et al. 2017). The main effect of fairness was significant, F (1, 30) = 237.07, $P < .001$, $\eta_p^2 = 0.89$, indicating that the participants made more punish amount when facing unfair distribution (2.39 ± 0.56) than fair distribution (0.71 ± 0.41). The main effect of empathy was significant, F (1, 30) = 86.33, $P < .001$, $\eta_p^2 = 0.74$, indicating that the participants made more punish amount in nonempathy condition (2.07 ± 0.49) than in empathy condition (1.04 ± 0.49). Importantly, the interaction between empathy and fairness was significant, F (1, 30) = 32.21, $P < 0.001$, $\eta_p^2 = 0.52$. Further simple effect analysis revealed that the punishment amount of nonempathy condition (0.93 ± 0.61) was higher than that of empathy (0.48 ± 0.51) for fair offers, F (1, 30) = 10.70, $P = .003$, $\eta_p^2 = 0.26$, and a greater difference between nonempathy (3.20 ± 0.74) and empathy (1.59 ± 0.70) condition for unfair offers, F (1, 30) = 97.86, $P < .001$, $\eta_p^2 = 0.77$ (see Fig. 2).

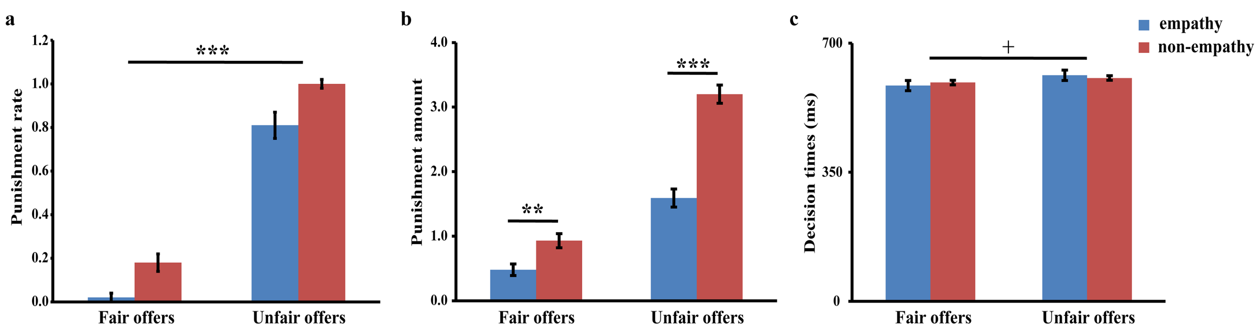

Figure 2. Bar graph of the punishment rate (a), the punishment amount (b) and decision times (c) in each condition. Error bars present standard error of means. *P < .10; **P < .01; ***P < .001.

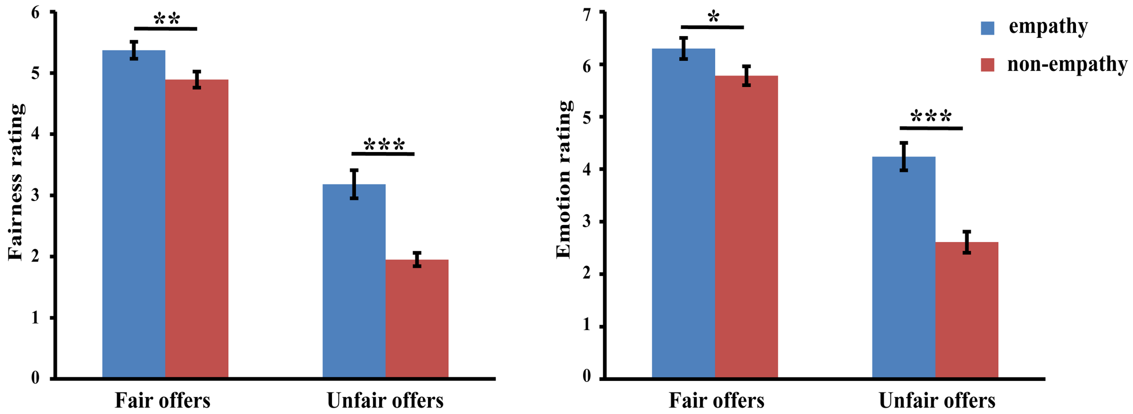

Figure 3. Bar graph of the fairness rating (left) and the emotion rating (right) in each condition. Error bars present standard error of means. *$P < .05$; **$P < .01$; ***$P < .001$.

Decision times

The main effect of fairness was marginally significant, F (1, 30) = 3.90, $P = .058$, $\eta_p^2 = 0.12$ (see Fig. 2), indicating that participants completed decisions more quickly for fair offers (588.85 ± 139.46 ms) than unfair (609.01 ± 135.79 ms). Neither other main effects nor interactions were significant (each $P > .10$).

Fairness perception

The main effect of fairness was significant, F (1, 30) = 259.12, $P < .001$, $\eta_p^2 = 0.90$, with higher fairness perception for fair offers (5.12 ± 0.64) than unfair (2.56 ± 0.71). The main effect of empathy was significant, F (1, 30) = 30.78, $P < 0.001$, $\eta_p^2 = 0.51$, with higher fairness perception for empathy condition (4.27 ± 0.86) than nonempathy (3.42 ± 0.39). The interaction effect of empathy × fairness was significant, F (1, 30) = 10.59, $P = .003$, $\eta_p^2 = 0.26$. Further simple effect analysis revealed that fairness perception was higher in empathy condition (5.37 ± 0.80) than in nonempathy condition (4.89 ± 0.70) for fair offers, F (1, 30) = 11.26, $P = .002$, $\eta_p^2 = 0.27$, and a greater difference between empathy (3.18 ± 1.24) and nonempathy (1.95 ± 0.54) condition for unfair offers, F (1, 30) = 28.13, $P < .001$, $\eta_p^2 = 0.48$ (see Fig. 3).

Emotion rating

The main effect of fairness was significant, F (1, 30) = 97.95, $P < 0.001$, $\eta_p^2 = 0.77$, with more positive emotion for fair offers (6.04 ± 0.94) than unfair offers (3.42 ± 0.96). The main effect of empathy was significant, F (1, 30) = 29.64, $P < .001$, $\eta_p^2 = 0.50$, with more positive emotion for empathy condition (5.27 ± 1.00) than nonempathy (4.20 ± 0.55). The interaction effect of empathy × fairness was significant, F (1, 30) = 11.07, $P = .002$, $\eta_p^2 = 0.27$. Further simple effect analysis revealed that emotion was more positive in empathy condition (6.30 ± 1.14) than in nonempathy condition (5.78 ± 1.01) for fair offers, F (1, 30) = 7.45, $P = .011$, $\eta_p^2 = 0.20$, and a greater difference between empathy (4.24 ± 1.46) and nonempathy (2.61 ± 1.10) condition for unfair offers, F (1, 30) = 27.36, $P < .001$, $\eta_p^2 = 0.48$ (see Fig. 3).

ERP results

AN1 (110–160 ms)

There was a significant interaction between empathy × fairness, F (1, 30) = 14.01, $P < .001$, $\eta_p^2 = 0.32$. Further simple effect analysis revealed that the fairness effect (Unfair minus Fair) was only significant for nonempathy condition, indicating a more negative AN1 in response to unfair offers than fair offers (−2.29 ± 2.00 μV vs. −1.26 ± 1.99 μV), F (1, 30) = 14.80, $P < .001$, $\eta_p^2 = 0.33$. However, the fairness effect on the AN1 was not significant for empathy condition (−1.57 ± 1.68 μV vs. −1.99 ± 2.38 μV), F (1, 30) = 1.43, $P = .241$. The empathy difference of this fairness effect is displayed in Fig. 4. The main effects of both empathy and fairness were not significant (Ps > .10).

MFN (290–330 ms)

A repeated-measures ANOVA of the MFN amplitude revealed a marginally significant main effect of empathy, F (1, 30) = 3.12, $P = .088$, $\eta_p^2 = 0.09$, indicating that the MFN was larger under nonempathy (−1.03 ± 3.35 μV) than empathy condition (−0.33 ± 3.12 μV). Furthermore, there was a significant effect of the interaction between empathy and fairness on the MFN amplitude, F (1, 30) = 5.16, $P = .031$, $\eta_p^2 = 0.15$. Consequently, a simple-effects analysis was conducted to investigate the significant interaction. Findings indicated that the MFN was more negative-going in response to unfair offers than fair offers in nonempathy condition (−1.53 ± 4.07 μV vs. −0.54 ± 2.92 μV), F (1, 30) = 5.65, $P = .024$, $\eta_p^2 = 0.16$. In contrast, the fairness effect on the MFN was not significant for empathy condition (−0.14 ± 3.29 μV vs. −0.52 ± 3.56 μV), F (1, 30) = 0.56, $P = .460$. The empathy difference of this fairness effect is displayed in Fig. 4. The main effect of fairness was not significant, F (1, 30) = 0.73, $P = .400$.

LPC (380–480 ms)

A repeated-measures ANOVA of the LPC amplitude revealed a significant main effect of empathy, F (1, 30) = 5.29, $P = .029$, $\eta_p^2 = 0.15$, indicating that the LPC was larger under empathy (2.93 ± 1.87 μV) than nonempathy condition (2.02 ± 2.45 μV). Furthermore, there was a marginally significant effect of the interaction between empathy and fairness on the LPC amplitude, F (1, 30) = 4.03, $P = .054$, $\eta_p^2 = 0.12$. Further simple effect analysis revealed that the fairness effect was only significant for nonempathy condition, indicating that LPC induced by fair offers (2.43 ± 2.41 μV) was larger than that of unfair offers (1.61 ± 2.86 μV) in nonempathy condition, F (1, 30) = 5.36, $P = .028$, $\eta_p^2 = 0.15$. However, there was no significant difference between fair (2.81 ± 2.16 μV) and unfair offers (3.06 ± 2.28 μV) in empathy condition, F (1, 30) = 0.31,

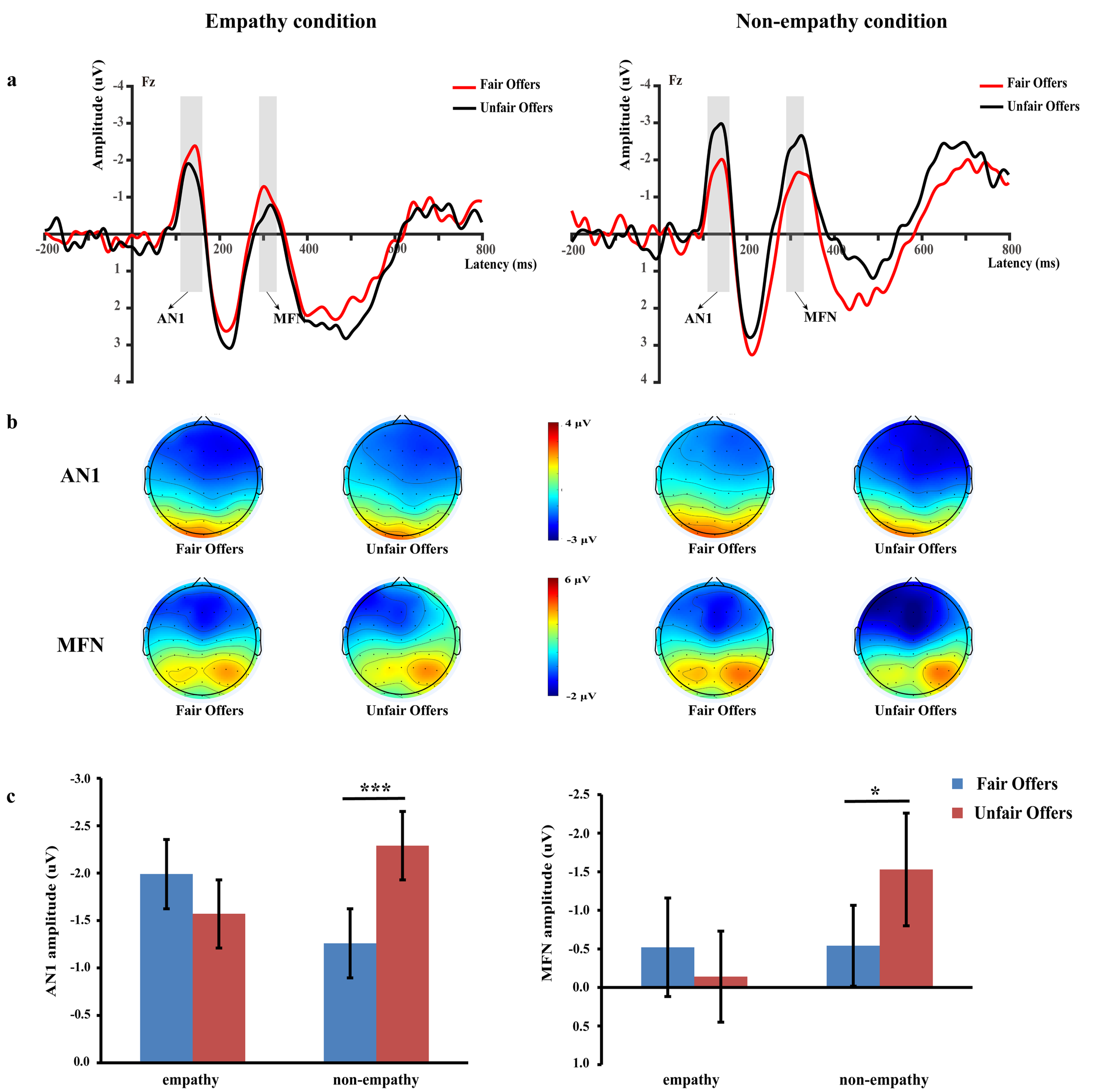

Figure 4. (a) Grand-average ERP waveforms at Fz. The grey bars highlight the time window of AN1 (110–160 ms) and MFN (290–330 ms); (b) Topographies voltage distribution of AN1 and MFN for each condition; (c) The bar graphs of mean AN1 and MFN values for each condition. Error bars present standard error of means. *$P < .05$, ***$P < .001$.

$P = .580$ (see Fig. 5.). The main effect of fairness was not significant, F (1, 30) = 0.98, $P = .331$.

Relationship between the ERP and behavioural data

To address the contribution of neural activity to behaviour in the corresponding conditions, we conducted a Spearman's correlation analysis between the behavioural data and the MFN components under four different conditions (Wang et al. 2022, Li et al. 2023b, Chen et al. 2024). For the emotion rating, significant positive correlations were obtained between emotion rating scores and mean MFN amplitudes under nonempathy condition regardless of fairness (unfair: $r = 0.385$, $P = .032$; fair: $r = 0.385$, $P = .033$) (see Fig. 6.). However, the correlation between emotion rating and the MFN for empathy condition was insignificant (unfair: $r = 0.199$, $P = .283$; fair: $r = -0.240$, $P = .193$). For the punishment rate, under the nonempathy condition, MFN was only negatively correlated with punishment rate ($r = -0.455$, $P = .010$) for the fair offers, while the correlation with the punishment rate of unfair offers is not significant ($r = -0.144$, $P = .438$) (see Fig. 7). And the correlation between punishment rate and the MFN for empathy condition was also insignificant (unfair: $r = -0.141$, $P = .448$; fair: $r = 0.024$, $P = .898$).

Discussion

This study used the TPP paradigm to investigate the neurophysiological effects of empathy on TPP and the underlying brain

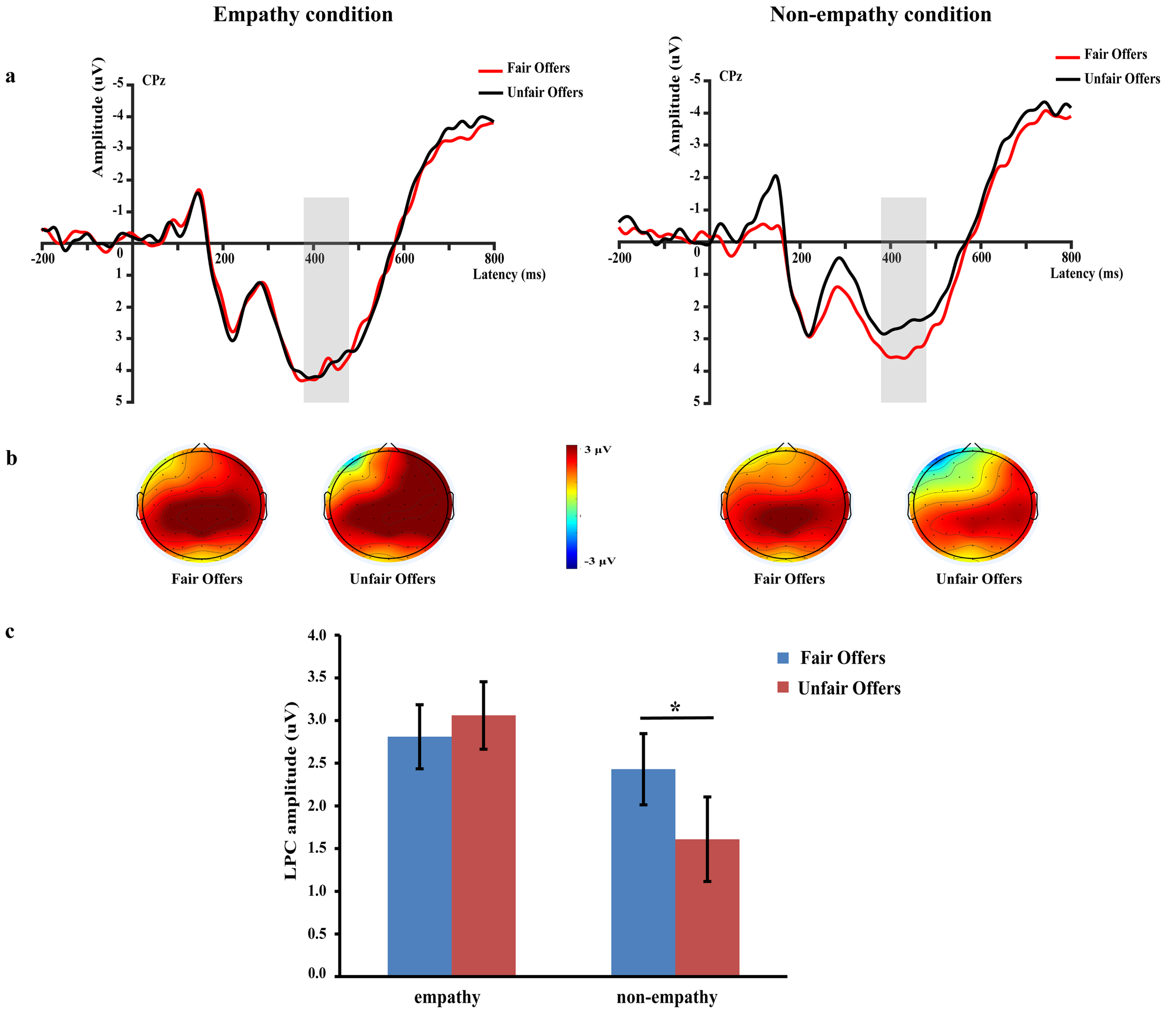

Figure 5. (a) Grand-average ERP waveforms at CPz. The grey bars highlight the time window of LPC (380–480 ms); (b) Topographies Voltage distribution of LPC for each condition; (c) The bar graphs of mean LPC values for each condition. Error bars present standard error of means. *$P < .05$.

activity and processes over time. The behavioural results showed that empathy and fairness interactively affected individual punishment amounts and subjective ratings. The ERP results revealed that empathy and fairness interacted to influence the early, middle, and late processing stages of TPP, as demonstrated by the predicted changes seen in the AN1, MFN, and LPC amplitudes.

Punishment amount, perception of fairness, and emotional experience are modulated by empathy and fairness

From a behavioural perspective, we used the punishment rate and punishment amount as indicators for individuals to impose TPP (Sun et al. 2015, Chen et al. 2024, Zheng, et al., 2024, Gao et al. 2025). The punishment rate represents an individual's decision to impose punishment on the 'perpetrator' or not. The punishment amount represents the intensity of the individual's decision to impose punishment on the 'perpetrator'. First, this study did not observe the interaction between empathy and fairness in punishment rate, which was inconsistent with our predictions. In contrast, the present study found that regardless of whether the 'perpetrator' of unfair allocation behaviour was empathic or nonempathic, as long as third-party individuals observed unfair allocation, they chose to make more decisions to punish than in the fair allocation condition in the decision of whether to punish or not. Previous research has also found that when people as third parties observe their friends act as 'perpetrators' of unfair behaviour, and when the victims of the outcome are disinterested others, the third-party individuals perceive the outcome to be more morally unacceptable (Li et al. 2022). This also means that when an individual engages in behaviour that harms the interests of others for their own selfish gain, the individual is regarded as more morally unacceptable (Zhan et al. 2020), regardless of their identity. Based on ethical fairness norms, participants were more likely to be influenced by their motivation for fairness when facing unfair allocations. Therefore, in our research, regardless of the identity of the 'perpetrators' of an unfair allocation, third-party individuals regarded this unfair allocation behaviour as morally unacceptable. At this time, the motivation for fairness may dominate individual decision-making; therefore, punishment decisions are made based on whether to choose punishment.

Importantly, consistent with our predictions, the punishment intensity of TPP is indeed influenced by empathy, and further

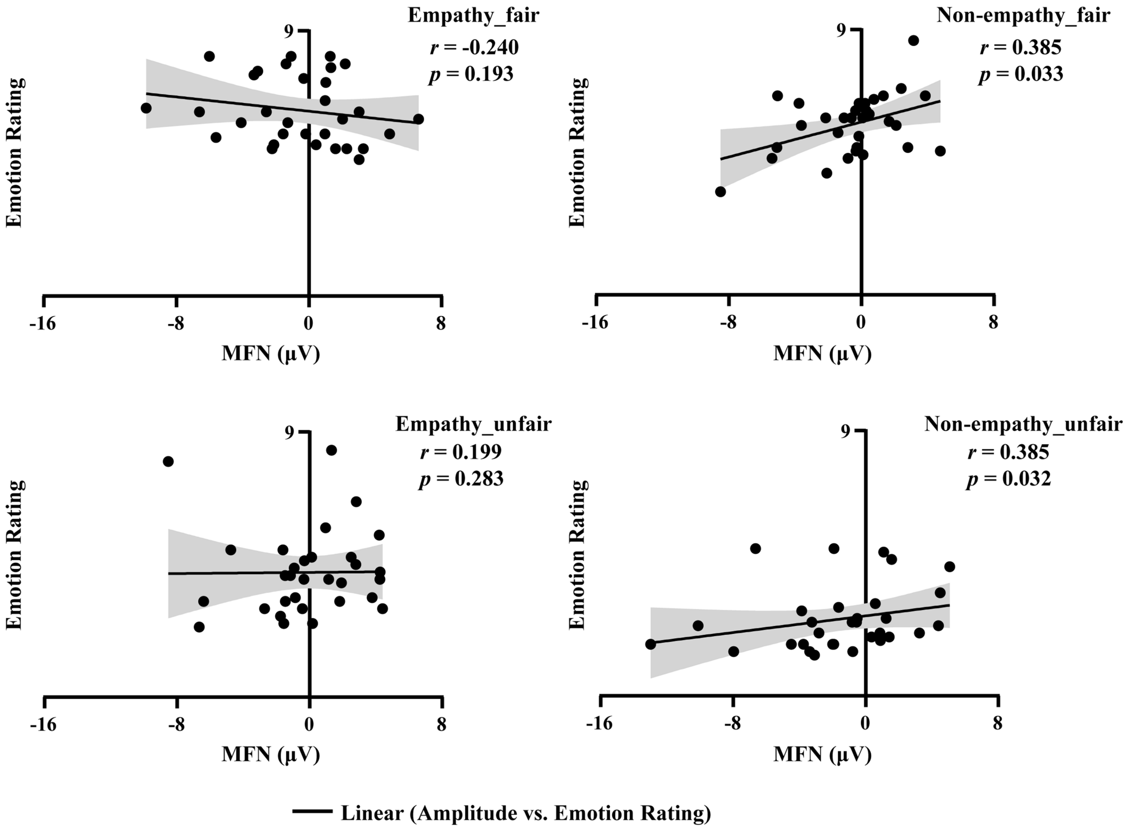

Figure 6. Spearman's correlations between MFN and emotion rating in four conditions, respectively. Note: The shaded grey region shows the 95% confidence intervals (CI).

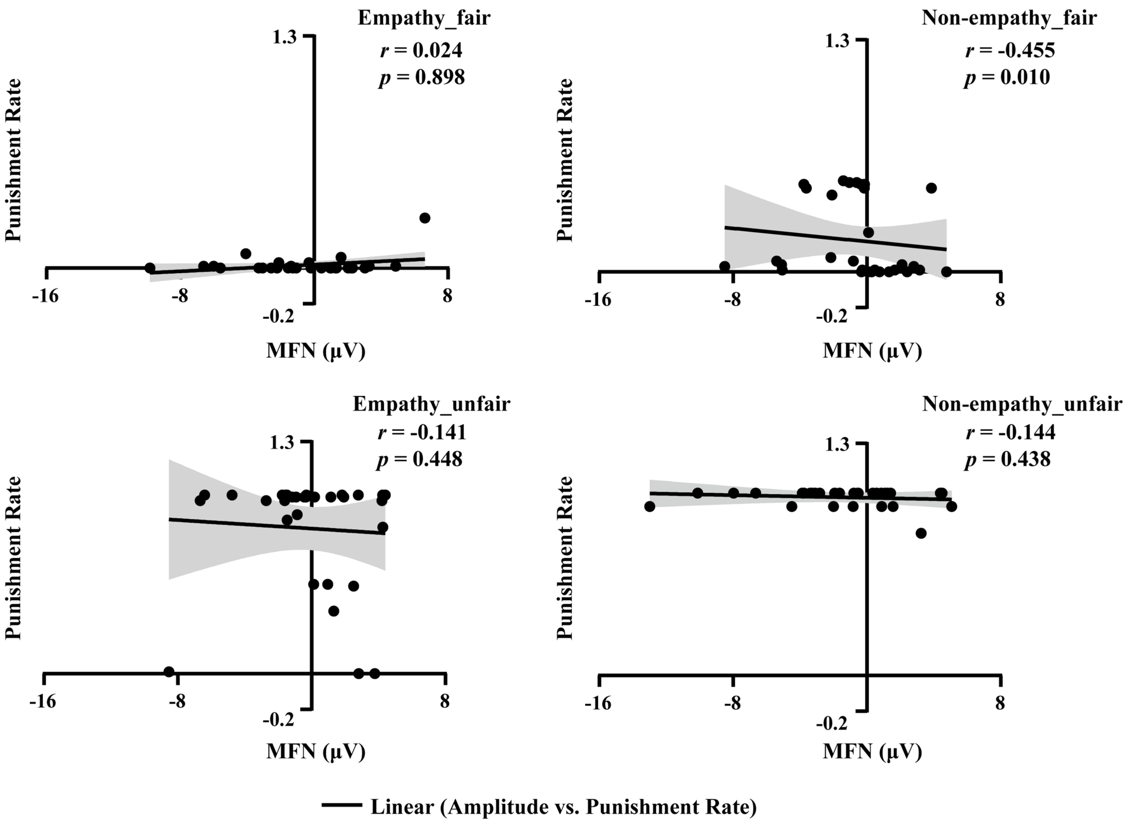

Figure 7. Spearman's correlations between MFN and punishment rate in four conditions, respectively. Note: The shaded grey region shows the 95% CI.

demonstrates the negative impact of empathetic altruism on fair decision-making: when faced with unfair allocations, participants imposed less punishment on empathetic 'perpetrators' than on non-empathetic 'perpetrators' and perceived their level of unfairness and negative emotions to be less intense. This is similar to previous studies in that, when an individual is a direct victim of unfair allocation, an empathic proposer allows the individual to have a higher tolerance of unfairness than a nonempathic proposer, which manifests in a higher acceptance rate of unfair allocation (He et al. 2022, Wang et al. 2022). This study also observed the interaction between empathy and fairness affects the punishment amount, which may be due to interference by empathy-altruistic motivation with fairness motivation when individuals face a trade-off between empathy altruistic motivation and fairness motivation. The details are as follows:

First, people prefer fairness in social life, and because of their inequality aversion, they tend to uphold fairness norms and impose TPPs when confronted with unfairness (Sun et al. 2015, Hu and Mai 2021). However, individuals' perceptions of fairness are not static or unchanging, but rather are dynamic and subject to the influence of other factors (Urbanska et al. 2019). For example, fairness perceptions are affected by identity, and people tend to believe that their own behaviour is perceived as fairer than that of others (Messick et al. 1985); unfair offers from in-group members are assessed as being fairer than unfair offers from out-group members (Wang et al. 2014, Guo et al. 2020, 2022). Similarly, individuals' perceived fairness is influenced by empathy processing related to unfair behaviour (Urbanska et al. 2019). For example, in the UG study, individuals had higher fairness perceptions and acceptance rates of unfair allocations from empathic identifiers than from nonempathic identifiers (He et al. 2022, Wang et al. 2022), which also implies that fairness perceptions are influenced by empathy. In this study, we also found that third-party decision makers evaluate unfair allocation behaviour as fairer when it is made by empathetic rather than non-empathetic proposers. Therefore, empathy influences individuals' fairness motives in TPP decisions.

Second, according to the empathy–altruism hypothesis, empathy enables people to empathize with others' pain, thereby generating an altruistic motivation to rescue others from suffering (Batson et al. 1991, 2015, Batson 2010). This means that in the process of empathy influencing third-party decision-making, there may involve a trade-off between empathy-altruistic motivation and fairness motivation, and that empathy-altruistic motivation interferes with fairness motivation, showing a tendency for empathy-altruistic motivation to be stronger than fairness motivation (He et al. 2022, Wang et al. 2022). This can be demonstrated by participants' empathy assessments of empathetic and non-empathetic proposers. In this study, we found that third-party observers showed more concern and emotional evaluation towards empathetic proposers than to non-empathetic proposers, indicating greater empathy for the pain experienced by empathetic proposers. At the same time, this further generated altruistic motives and manifested itself in participants' subjective assessments of the amount of punishment for empathic versus nonempathic proposers, as well as their emotions and perceptions of fairness. We found that third-party decision makers tended to report higher perceptions of fairness and more positive emotional experiences regarding unfair allocation behaviours by empathic compared to nonempathic unfair allocators, which led individuals to reduce the intensity of punishment for their behaviour. This result may have been caused by altruistic motivation after empathizing, where third-party decision makers believe that the unfair distribution behaviour of empathetic individuals occurs 'for a reason' (such as third-party decision makers believing that left-behind children's families need more financial help due to difficulties), and this reason is not a malicious behaviour that harms others (Treadway et al. 2014). This empathetic altruistic motivation reduces individuals' unfairness aversion and negative emotions (He et al. 2022), thereby reducing punishment intensity. In summary, the effect of empathy on TPP decisions may stem from interference by empathy-altruistic motives with fairness motives.

In summary, empathy has different impacts on the decision to punish (punishment rate) and the intensity of punishment (punishment amount) in TPP. As mentioned above, the decision to punish or not might be mainly influenced by the motive of fairness, whereas in terms of the intensity of punishment, empathy comes into play, and the altruistic motive interferes with the fairness motive, which in turn influences the fairness decision. This is common in real-life situations. When the public attempts to balance between law and reason, in terms of criminals who attempt to induce public empathy, although the public considers that 'extenuating circumstances' exist, they still believe that they should be punished, although the extent of the punishment can be mitigated. In short, empathy increased third-party tolerance towards the unfair offers of empathetic allocators, which was reflected in the degree of punishment rather than the decision regarding whether to punish or not; at the same time, the altruistic motivation of individuals to value the welfare of empathetic allocators exceeded their original fairness motivation, resulting in third-party individuals showing 'favoritism' towards empathetic allocators.

The AN1, MFN, and LPC are modulated by empathy and fairness

The ERP results further confirmed these findings. According to previous ERP research, fairness decision-making mainly consists of three distinct stages: early rapid and automated intuitive fairness processing, medium-term immediate and intense emotional responses, and late slow but elaborative cognitive processes (Wang et al. 2014, Hu and Mai 2021, He et al. 2022, Chen et al. 2024).

Empathy modulates the early stage of TPP processing

The early stages of fairness processing in TPP are represented by AN1, which reflects the initial automated processing of decisions (Wang et al. 2014, He et al. 2022). AN1 is associated with early automated attention, and stimuli that defy expectations are generally more likely to attract attention (Boudreau et al. 2009, Annic et al. 2014). Research has found that larger AN1 wave amplitudes reflect an individual's perceived degree of expectation violation towards different allocation proposals (Wang et al. 2014). In this study, the fairness effect of AN1 was modulated by empathy; in TPP, the fairness effect was only meaningful for proposers with nonempathic identities, and third-party observers had larger AN1 amplitudes for unfair than for fair offers. This is consistent with our hypothesis and previous research results (Wang et al. 2014, He et al. 2022), which may be due to whether the proposal meets expectations. People engage in psychological theory reasoning and processing in social decision-making, forming predictions of others' behaviour (Wu and Zhou 2012, Schurz et al. 2021, He et al. 2022), and these predictions subsequently influence people's decision-making (Pruitt and Kimmel 1977). Numerous studies have shown that people follow fairness norms in social life and express an aversion to unfairness (Sun et al. 2015, Hu and Mai 2021); as a result, they form expectations of fair allocation from others and view unfair allocation as a stimulus that violates expectations. Therefore, under nonempathy conditions, third-party observers quickly grasp the presence of unfairness and the AN1 exhibits sensitivity to fairness effects. As shown by the behavioural outcomes, altruistic motivation for empathy can affect people's perceptions of fairness. Therefore, third-party decision makers are more able to empathize with the difficulties they face when forming expectations concerning empathetic proposers and believe that their unfair allocation is reasonable and justifiable (He et al. 2022, Wang et al. 2022), which shapes the expectation that both fair and unfair allocations are acceptable. That is, the altruistic motivation generated by empathy reduces people's inequity aversion and weakens their perception of fairness effects; therefore, AN1 is not sensitive to fairness effects under the empathy condition.

Empathy modulates the middle stage of TPP processing

The middle stage of fairness processing in TPP is represented by MFN, which reflects semiautomatic processing of decision-making (Wang et al. 2014, He et al. 2022). Research has found that larger MFN amplitudes can reflect deviations from individual expectations, individuals' aversion to violating fairness norms, and the resulting cognitive and emotional conflicts in the decision-making process (Spape et al. 2019, He et al. 2022, Zinchenko et al. 2023). In the present study, the fairness effect of MFN was also modulated by empathy, and this effect was only meaningful for proposers with nonempathic identities, with third-party observers having greater MFN amplitudes for unfair than for fair offers. This is consistent with our hypothesis and previous research results (Wang et al. 2022), which may be due to the negative emotions of unfair aversion caused by expected violations or violations of fair norms.

On the one hand, according to the expectancy-deviation hypothesis of MFN (Oliveira et al. 2007, Hewig et al. 2011), MFN originates from the anterior cingulate cortex region, which is responsible for encoding the deviation between an actual outcome and expectations, and the larger this deviation is, the larger the induced MFN amplitude (Si et al. 2019, Yang et al. 2024). As mentioned previously, the principle of fairness is pervasive in human societies; thus, people expect others to make fair allocations (Wang et al. 2014). Simultaneously, the prominence of empathic identities affects an individual's assessment of fairness, which, in turn, alters the individual's expectation that extenuating circumstances exist regarding unfair allocations by an empathic proposer (Wang et al. 2022). Therefore, the fairness effect of MFN was reflected in nonempathy rather than in empathy conditions.

On the other hand, MFN is also related to violations of social norms, which can trigger larger MFN amplitudes (Zinchenko et al. 2023). Given that MFN is more pronounced following negative stimuli (Paul and Pourtois 2017), this study found a significant negative correlation between MFN and punishment rate under the nonempathy-fair condition, which also means that larger MFN amplitude predicts higher punishment rate (Chen et al. 2024). The emotional model suggests that negative emotions drive individuals to engage in TPP (Fehr and Gächter 2002, Xiao and Houser 2005, Zheng, et al., 2024). Violations of fairness norms can create unfairness aversion and negative emotional experiences (Sun et al. 2015, Hu and Mai 2021), which reflect the fairness effect of MFN that was significant in the nonempathy condition. Correlation analysis between MFN amplitudes and emotional assessment also confirmed this point. The results showed that under non-empathetic conditions, MFN amplitude was significantly positively correlated with emotional assessment. This also means that under nonempathy conditions, the larger the MFN amplitude, the stronger the negative emotional assessment. This effect was not observed under empathy conditions, which may indicate that empathy reduces aversion to unfairness and the negative emotions it triggers. Therefore, the fairness effect of MFN appears only under nonempathy rather than empathy conditions.

Empathy modulates the late stage of TPP processing

The LPC represents the late stage of fairness processing in TPP, which reflects a detailed decision-making process (Cui et al. 2019, Hu and Mai 2021). Research has shown that the LPC is associated with cognitive resource investment and advanced motivation and emotion assessment. As cognitive resource investment and motivation increase, the amplitude of the LPC decreases (Cui et al. 2019, Zheng, et al., 2024, Zheng, et al., 2024;). In this study, the fairness effect of the LPC was similar to that of early to middle stage processing, and was also modulated by empathy. Specifically, in the nonempathy condition, unfair offers elicited greater LPC amplitudes than fair offers; this difference disappeared in the empathy condition. This finding is consistent with our hypotheses and with previous research results (Cui et al. 2019, Zheng, et al., 2024). The decision to punish in TPP results from a trade-off between an individual's negative emotions and self-interest mechanisms (Zheng, et al., 2024), and this trade-off process requires individuals to invest more cognitive resources in exhaustive consideration. The relationship between this trade-off and cognitive resources has also been confirmed by the results of decision-making time (Zhan et al. 2020), which manifested as third-party individuals needing longer to make decisions when faced with unfair rather than fair allocations. Under nonempathy conditions, on the one hand, driven by negative emotions induced by unfair aversion, third-party observers attempt to make punitive decisions; on the other hand, this type of punishment requires individuals to pay a cost, and this self-interest motive hinders individuals from making punishment decisions. This dilemma requires individuals to invest cognitive resources in the evaluation, which is reflected in the fairness effect of the LPC (Zheng, et al., 2024). In contrast, in the empathy condition, third-party observers were more tolerant of empathic violators (Ouyang et al. 2021, He et al. 2022, Wang et al. 2022), which, as shown by the fairness effect of the MFN, did not produce a difference in negative affect, and thus did not involve such a trade-off. Therefore, we speculated that due to this trade-off, third-party observers invested more cognitive resources in deciding whether to punish the violator, and that this discrepancy appeared in the nonempathy rather than empathy condition.

In summary, this study provides a new perspective for investigating the impact of empathy on TPP. Our results suggest that unfair behaviour without empathy for the 'perpetrator' induces strong unfairness aversion and negative emotions, as well as stronger trade-offs, which are reflected in significant AN1, MFN, and LPC fairness effects in punishment decisions for nonempathic proposers. However, this influence disappears when a third-party empathizes with the 'perpetrator', as empathy-altruistic motives interfere with fairness motives and increase tolerance for unfair behaviour. In summary, the results of this study reveal the potential psychological and neural mechanisms between altruistic and fairness motives in the decision-making process of TPP.

Implications and future directions

The findings of the present study have several implications. Previous research has focused on the fact that inducing individuals to empathize with unjust actors increases the tolerance of second-party individuals. The present study found that inducing third-party individuals to empathize with unjust actors does the same, although this tolerance is reflected in the degree of punishment. This finding broadens previous insights into the psychological and neural mechanisms of empathy in fairness decision-making. Practically, this study provides insights into judicial system decisions and social news incidents. For example, if the perpetrator inflicts unfair behaviour on the victim, the perpetrator may use media exaggeration to induce empathy among the public or sentencing personnel to evade or mitigate punishment. The phenomenon of interference by empathy-altruistic motives with fairness motives reminds the public of the need to rationally view the phenomenon of 'empathy' in society: to distinguish the true intention of the 'perpetrator' to harm, rather than blindly empathizing with the harm they have suffered.

This study has some limitations that should be considered. First, this study used only Chinese participants; however, TPP could be influenced by differences between Chinese people and Westerners (Guo et al. 2022), and subsequent studies could consider different groups of participants to broaden generalizability. Second, the emotional assessment in this study did not explore in detail which negative emotions were at play. For example, previous studies have found that both the negative emotions of disgust and anger can play a role (Ginther et al. 2022, Fan et al. 2024); thus, subsequent studies could explore the role of emotions in detail. Finally, TPP was conducted in the laboratory, and it is necessary to explore it further using a more ecologically valid adaptation paradigm in the future.

Conclusion

In summary, this study used an ERP approach to reveal the roles and neural mechanisms of empathy in TPP. Third-party individuals increased their tolerance for unfair offers of empathetic allocators because of empathy-altruistic motives, which manifested as a reduction in rather than an exemption from punishment. In addition, empathy for the 'perpetrator' reduces the individual's sensitivity to fairness: when there is no empathy for the 'perpetrator', individuals have stronger negative emotions in response to the unfair allocation, as well as a stronger motivation and investment of cognitive resources in punishing the unfair behaviour; this discrepancy disappeared in the empathy condition. In summary, the results of this study reveal the psychological and neural mechanisms underlying the relationship between empathy-altruistic and fairness motives in third-party decision-making processes.

Supplementary data

Supplementary data is available at *SCAN* online.

Conflict of interest: There are no conflicts of interest to declare.

Funding

This research was supported by the National Natural Science Foundation of China (grant number: 32471124) and the Natural Science Foundation of Hunan Province [grant number: 2023JJ30400 and 2025JJ50160].

Data availability statement

All data supporting the findings of this study are available from the corresponding author on reasonable request.

Ethics approval statement

This study was conducted in accordance with the Declaration of Helsinki. This study was approved by the Research Ethics Committee of the Department of Psychology at Hunan Normal University, China (Ethics Approval Ref. No. 2023-287).

References

Annic A, Bocquillon P, Bourriez JL *et al*. Effects of stimulus-driven and goal-directed attention on prepulse inhibition of the cortical responses to an auditory pulse. *Clin Neurophysiol* 2014;**125**:1576–88. https://doi.org/10.1016/j.clinph.2013.12.002

Batson CD. Empathy-induced altruistic motivation. In: Mikulincer M and Shaver PR (eds), *Prosocial motives, emotions, and behavior: the better angels of our nature*. Washington, DC, US: American Psychological Association, 2010, 15–34

Batson CD, Batson JG, Slingsby JK *et al*. Empathic joy and the empathy-altruism hypothesis. *J Pers Soc Psychol* 1991;**61**:413–26. https://doi.org/10.1037/0022-3514.61.3.413

Batson CD, Klein TR, Highberger L *et al*. Immorality from empathy-induced altruism: when compassion and justice conflict. *J Pers Soc Psychol* 1995;**68**:1042–54. https://doi.org/10.1037/0022-3514.68.6.1042

Batson CD, Lishner DA, Stocks EL. The empathy-altruism hypothesis. In: Schroeder DA and Graziano WG (eds), *The oxford handbook of prosocial behavior*. New York, NY, US: Oxford University Press, 2015, 259–81. https://doi.org/10.1093/oxfordhb/9780195399813.013.023

Batson CD, Shaw LL. Evidence for Altruism: toward a pluralism of prosocial motives. *Psychol Inq* 1991;**2**:107–22. https://doi.org/10.1207/s15327965pli0202_1

Boudreau C, McCubbins MD, Coulson S. Knowing when to trust others: an ERP study of decision making after receiving information from unknown people. *Soc Cogn Affect Neurosci* 2009;**4**:23–34. https://doi.org/10.1093/scan/nsn034

Bouwer FL, Honing H, Slagter HA. Beat-based and memory-based temporal expectations in rhythm: similar perceptual effects, different underlying mechanisms. *J Cognitive Neurosci* 2020;**32**:1221–41. https://doi.org/10.1162/jocn_a_01529

Bradley MM, Codispoti M, Cuthbert BN *et al*. Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion* 2001;**1**:276–98. https://doi.org/10.1037//1528-3542.1.3.276

Chen P, Liu S, Zhang Y *et al*. Same allocation proposed by an individual or a group elicits distinct responses: evidence from event-related potentials and neural oscillation. *Neuroimage* 2024;**290**:120565. https://doi.org/10.1016/j.neuroimage.2024.120565

Cui F, Wang C, Cao Q *et al*. Social hierarchies in third-party punishment: a behavioral and ERP study. *Biol Psychol* 2019;**146**:107722. https://doi.org/10.1016/j.biopsycho.2019.107722

Decety J. The neural pathways, development and functions of empathy. *Curr Opin Behav Sci* 2015;**3**:1–6. https://doi.org/10.1016/j.cobeha.2014.12.001

Decety J, Bartal IB-A, Uzefovsky F *et al*. Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philos Trans R Soc London, Ser B* 2016;**371**:20150077. https://doi.org/10.1098/rstb.2015.0077

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 2004;**134**:9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Ding I, Javed J. The autocrat's moral-legal dilemma: popular morality and legal institutions in China. *Comp Political Stud* 2020;**54**:989–1022. https://doi.org/10.1177/0010414020957694

Duzcu H, Özkurt TE, Mapelli I *et al*. N1-P2: neural markers of temporal expectation and response discrimination in interval timing. *Acta Neurobiol. Exp* 2019;**79**:193–204. https://doi.org/10.21307/ane-2019-0017

Fan L, Molho C, Kupfer TR *et al*. Moral violations that target more valued victims elicit more anger, but not necessarily more disgust. *J Exp Soc Psychol* 2024;**112**:104597. https://doi.org/10.1016/j.jesp.2024.104597

Faul F, Erdfelder E, Lang A-G *et al*. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007;**39**:175–91. https://doi.org/10.3758/BF03193146

Fehr E, Fischbacher U. Third-party punishment and social norms. *Evol Hum Behav* 2004;**25**:63–87. https://doi.org/10.1016/s1090-5138(04)00005-4

Fehr E, Gächter S. Altruistic punishment in humans. *Nature* 2002;**415**:137–40. https://doi.org/10.1038/415137a

Gao Y, Ao L, Wang H *et al*. Render help or stand by? The effect of group size on third-party punishment and its neural mechanisms. *Behav Brain Res*. 2025;**476**:115256. https://doi.org/10.1016/j.bbr.2024.115256

Gehring WJ, Willoughby AR. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science (New York, N Y)* 2002;**295**:2279–82. https://doi.org/10.1126/science.1066893

Ginther MR, Hartsough LES, Marois R. Moral outrage drives the interaction of harm and culpable intent in third-party punishment decisions. *Emotion* 2022;**22**:795–804. https://doi.org/10.1037/emo0000950

Gu RL, He Y, Cui F. Reconsidering the relationship between empathy and prosocial behavior. *Chin Sci Bull*. 2025;**70**:982–990. https://doi.org/10.1360/TB-2024-0025

Gummerum M, Hanoch Y, Keller M *et al*. Preschoolers' allocations in the dictator game: the role of moral emotions. *Journal of Economic Psychology* 2010;**31**:25–34. https://doi.org/10.1016/j.joep.2009.09.002

Guo R, Ding J, Wu Z. How intergroup relation moderates group bias in third-party punishment. *Acta Psychologica* 2020;**205**:103055. https://doi.org/10.1016/j.actpsy.2020.103055

Guo Z, Guo R, Xu C *et al*. Reflexive or reflective? Group bias in third-party punishment in Chinese and Western cultures. *J Exp Soc Psychol* 2022;**100**:104284. https://doi.org/10.1016/j.jesp.2022.104284

Güth W, Schmittberger R, Schwarze B. An experimental analysis of ultimatum bargaining. *J Econ Behav Organiz* 1982;**3**:367–88. https://doi.org/10.1016/0167-2681(82)90011-7

Haegerich TM, Bottoms BL. Empathy and jurors' decisions in patricide trials involving child sexual assault allegations. *Law Hum Behav* 2000;**24**:421–48. https://doi.org/10.1023/A:1005592213294

He YJ, Hu XM, Mai XQ. Influence of empathic concern on fairness-related decision making: evidence from ERP. *Acta Psychol Sin* 2022;**54**:385–97. https://doi.org/10.3724/sp.J.1041.2022.00385

Hernandez Pena L, Weidacker K, Massau C *et al*. Can people empathize with offenders and victims during violent scenes? Behavioral and brain correlates of affective and cognitive empathy considering victim vs. offender perspective using the Bochumer affective and cognitive empathy task (BACET). *Neuropsychologia* 2024;**194**:108784. https://doi.org/10.1016/j.neuropsychologia.2023.108784

Hewig J, Kretschmer N, Trippe RH *et al*. Why humans deviate from rational choice. *Psychophysiol* 2011;**48**:507–14. https://doi.org/10.1111/j.1469-8986.2010.01081.x

Hoffman ML. Empathy and prosocial behavior. In: Lewis M, Haviland-Jones JM, Barrett LF (eds), *Handbook of Emotions*. 3rd edn New York: The Guilford Press, 2008, 440–55.

Holroyd CB, Krigolson OE. Reward prediction error signals associated with a modified time estimation task. *Psychophysiol* 2007;**44**:913–17. https://doi.org/10.1111/j.1469-8986.2007.00561.x

Hu X, Mai X. Social value orientation modulates fairness processing during social decision-making: evidence from behavior and brain potentials. *Soc Cogn Affect Neurosci* 2021;**16**:670–82. https://doi.org/10.1093/scan/nsab032

Hu X, Zhang Y, Liu X *et al*. Effects of transcranial direct current stimulation over the right dorsolateral prefrontal cortex on fairness-related decision-making. *Soc Cogn Affect Neurosci* 2022;**17**:695–702. https://doi.org/10.1093/scan/nsac004

Jaroslawska AJ, McCormack T, Burns P *et al*. Outcomes versus intentions in fairness-related decision making: school-aged children's decisions are just like those of adults. *J Exp Child Psychol*. 2020;**189**:104704. https://doi.org/10.1016/j.jecp.2019.104704

Johnson R, Barnhardt J, Zhu J. The deceptive response: effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biol Psychol* 2003;**64**:217–53. https://doi.org/10.1016/j.biopsycho.2003.07.006

Johnson R, Henkell H, Simon E *et al*. The self in conflict: the role of executive processes during truthful and deceptive responses about attitudes. *Neuroimage* 2008;**39**:469–82. https://doi.org/10.1016/j.neuroimage.2007.08.032

Jordan JJ, Hoffman M, Bloom P *et al*. Third-party punishment as a costly signal of trustworthiness. *Nature* 2016;**530**:473–76. https://doi.org/10.1038/nature16981

Kundro TG, Nurmohamed S, Kakkar H *et al*. Time and punishment: time delays exacerbate the severity of third-party punishment. *Psychol Sci* 2023;**34**:914–31. https://doi.org/10.1177/09567976231173900

Leliveld MC, Dijk E, Beest I. Punishing and compensating others at your own expense: the role of empathic concern on reactions to distributive injustice. *Eur J Social Psychol* 2012;**42**:135–40. https://doi.org/10.1002/ejsp.872

Li J, Li M, Sun Y *et al*. Interpersonal relationships modulate subjective ratings and electrophysiological responses of moral evaluations. *Cogn Affect Behav. Neurosci* 2022;**23**:125–41. https://doi.org/10.3758/s13415-022-01041-9

Li M, Li J, Zhang G *et al*. Social distance modulates the influence of social observation on pro-environmental behavior: an event-related potential (ERP) study. *Biol Psychol* 2023b;**178**:108519. https://doi.org/10.1016/j.biopsycho.2023.108519

Liu X, Hu X, Shi K *et al*. Your losses are mine: the influence of empathic concern on evaluative processing of others' outcomes. *Cogn Affect Behav Neurosci* 2020;**20**:481–92. https://doi.org/10.3758/s13415-020-00779-4

Liu Y, Bian X, Hu Y *et al*. Intergroup bias influences third-party punishment and compensation: in-group relationships attenuate altruistic punishment. *Soc Behav Pers* 2018;**46**:1397–408. https://doi.org/10.2224/sbp.7193

Liu Y, Li L, Zheng L *et al*. Punish the perpetrator or compensate the victim? Gain vs. loss context modulate third-party altruistic behaviors. *Front Psychol* 2017;**8**:2066. https://doi.org/10.3389/fpsyg.2017.02066

Luck SJ. Multiple mechanisms of visual-spatial attention: recent evidence from human electrophysiology. *Behav Brain Res*. 1995;**71**:113–23. https://doi.org/10.1016/0166-4328(95)00041-0

Luo L, Xu H, Tian X *et al*. The neurocomputational mechanism underlying decision-making on unfairness to self and others. *Neurosci Bull* 2024;**40**:1471–88. https://doi.org/10.1007/s12264-024-01245-8

Messick DM, Bloom S, Boldizar JP *et al*. Why we are fairer than others. *J Exp Soc Psychol* 1985;**21**:480–500. https://doi.org/10.1016/0022-1031(85)90031-9

Mothes H, Enge S, Strobel A. The interplay between feedback-related negativity and individual differences in altruistic punishment: an EEG study. *Cogn Affect Behav Neurosci* 2016;**16**:276–88. https://doi.org/10.3758/s13415-015-0388-x

Oliveira FT, McDonald JJ, Goodman D. Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J Cognitive Neurosci* 2007;**19**:1994–2004. https://doi.org/10.1162/jocn.2007.19.12.1994

Ouyang H, Yu J, Duan J *et al*. Empathy-based tolerance towards poor norm violators in third-party punishment. *Exp Brain Res* 2021;**239**:2171–80. https://doi.org/10.1007/s00221-021-06128-2

Paul K, Pourtois G. Mood congruent tuning of reward expectation in positive mood: evidence from FRN and theta modulations. *Soc Cogn Affect Neurosci* 2017;**12**:765–74. https://doi.org/10.1093/scan/nsx010

Peters SL, van den Bos K. When fairness is especially important: reactions to being inequitably paid in communal relationships. *Soc Justice Res* 2008;**21**:86–105. https://doi.org/10.1007/s11211-007-0056-6

Pfattheicher S, Sassenrath C, Keller J. Compassion magnifies third-party punishment. *J Pers Soc Psychol* 2019;**117**:124–41. https://doi.org/10.1037/pspi0000165

Picó A, Gračanin A, Gadea M *et al*. How visible tears affect observers' judgements and behavioral intentions: sincerity, remorse, and punishment. *J Nonverbal Behav* 2020;**44**:215–32. https://doi.org/10.1007/s10919-019-00328-9

Pillutla MM, Murnighan JK. Unfairness, anger, and spite: emotional rejections of ultimatum offers. *Organ Behav Hum Decis Process* 1996;**68**:208–24. https://doi.org/10.1006/obhd.1996.0100

Plöchl M, Ossandón JP, König P. Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Front Human Neurosci* 2012;**6**:278. https://doi.org/10.3389/fnhum.2012.00278

Pruitt DG, Kimmel MJ. Twenty years of experimental gaming: critique, synthesis, and suggestions for the future. *Annu Rev Psychol* 1977;**28**:363–92. https://doi.org/10.1146/annurev.ps.28.020177.002051

Schurz M, Radua J, Tholen MG *et al*. Toward a hierarchical model of social cognition: a neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychol Bull* 2021;**147**:293–327. https://doi.org/10.1037/bul0000303

Si Y, Wu X, Li F *et al*. Different decision-making responses occupy different brain networks for information processing: a study based on EEG and TMS. *Cereb. Cortex* 2019;**29**:4119–29. https://doi.org/10.1093/cercor/bhy294

Spape M, Harjunen V, Ahmed I *et al*. The semiotics of the message and the messenger: how nonverbal communication affects fairness perception. *Cogn Affect Behav Neurosci* 2019;**19**:1259–72. https://doi.org/10.3758/s13415-019-00738-8

Sun L, Tan P, Cheng Y *et al*. The effect of altruistic tendency on fairness in third-party punishment. *Front Psychol* 2015;**6**:820. https://doi.org/10.3389/fpsyg.2015.00820

Treadway MT, Buckholtz JW, Martin JW *et al*. Corticolimbic gating of emotion-driven punishment. *Nat Neurosci* 2014;**17**:1270–75. https://doi.org/10.1038/nn.3781

Tsoudis O. The influence of empathy in mock jury criminal cases: adding to the affect control model. *West Criminol Rev* 2002;**4**:55–67.

Urbanska K, McKeown S, Taylor LK. From injustice to action: the role of empathy and perceived fairness to address inequality via victim compensation. *J Exp Soc. Psychol* 2019;**82**:129–40. https://doi.org/10.1016/j.jesp.2019.01.010

Wang A, Zhu L, Lyu D *et al*. You are excusable! Neural correlates of economic neediness on empathic concern and fairness perception. *Cogn. Affect Behav Neurosci* 2022;**22**:99–111. https://doi.org/10.3758/s13415-021-00934-5

Wang G, Li J, Li Z *et al*. Medial frontal negativity reflects advantageous inequality aversion of proposers in the ultimatum game: an ERP study. *Brain Res*. 2016;**1639**:38–46. https://doi.org/10.1016/j.brainres.2016.02.040

Wang YW, Zhang Z, Zhang W *et al*. Group membership modulates recipient's fairness consideration in ultimatum game. *Acta Psychol. Sin* 2014;**46**:1850–59. https://doi.org/10.3724/sp.J.1041.2014.01850

Wu Y, Zhou XL. The context-dependency of fairness processing: evidence from ERP study. *Acta Psychol. Sin* 2012;**44**:797–806. https://doi.org/10.3724/sp.J.1041.2012.00797

Xiao E, Houser D. Emotion expression in human punishment behavior. *Proc Natl Acad Sci* 2005;**102**:7398–401. https://doi.org/10.1073/pnas.0502399102

Yang J, Gu R, Liu J *et al*. To blame or not? Modulating third-party punishment with the framing effect. *Neurosci Bull* 2022;**38**:533–47. https://doi.org/10.1007/s12264-021-00808-3

Yang L, Gao Y, Ao L *et al*. Context modulates perceived fairness in altruistic punishment: neural signatures from ERPs and EEG oscillations. *Brain Topography* 2024;**37**:764–82. https://doi.org/10.1007/s10548-024-01039-1

Yin XL, Li JB, Chen SY *et al*. Neural mechanisms of third-party punishment: evidence from transcranial direct current stimulation. *Acta Psychol. Sin* 2019;**51**:571–83. https://doi.org/10.3724/sp.J.1041.2019.00571

Zhan Y, Xiao X, Tan Q *et al*. Neural correlations of the influence of self-relevance on moral decision-making involving a trade-off between harm and reward. *Psychophysiol* 2020;**57**:e13590. https://doi.org/10.1111/psyp.13590

Zhang Q, Deng NL, Jiang XM *et al*. The time course of self-relevance affecting emotional word processing. *Acta Psychol. Sin* 2020;**52**: 946–957 https://doi.org/10.3724/sp.J.1041.2020.00946

Zheng H, Chen R, Mai X The cognitive and neural mechanism of third-party punishment. *Adv Psychol Sci*. 2024;**32**: 398–412 https://doi.org/10.3724/SP.J.1042.2024.00398

Zheng L, Zhou Y, Ouyang H *et al*. Receivers' responses are integrated into costly third-party punishment in a way that interacts with the unfairness of allocations. *Brain Res Bull* 2024;**217**:111082. https://doi.org/10.1016/j.brainresbull.2024.111082

Zinchenko O, Gorin A, Revazyan A *et al*. Electrophysiological correlates of third-party punishment: ERP study. *Neurosci Lett* 2023;**808**:137276. https://doi.org/10.1016/j.neulet.2023.137276

Social Cognitive and Affective Neuroscience, 2025, 20(1), nsaf042, DOI: https://doi.org/10.1093/scan/nsaf042, Advance Access Publication Date: 2 May 2025, Original Research – Neuroscience
Received: 29 October 2024; **Revised**: 28 March 2025; **Accepted**: 29 April 2025