

RESEARCH ARTICLE

Reinforcement learning of altruistic punishment differs between cultures and across the lifespan

Ziyan Guo^{1,2}, Jialu Yu^{1,2}, Wenxin Wang^{1,2}, Patricia Lockwood³, Zhen Wu^{1,2*}

1 Department of Psychological and Cognitive Sciences, Tsinghua University, Beijing, China, **2** Lab for Lifelong Learning, Tsinghua University, Beijing, China, **3** Centre for Human Brain Health and Institute for Mental Health, School of Psychology, University of Birmingham, Birmingham, United Kingdom

* zhen-wu@tsinghua.edu.cn**OPEN ACCESS**

Citation: Guo Z, Yu J, Wang W, Lockwood P, Wu Z (2024) Reinforcement learning of altruistic punishment differs between cultures and across the lifespan. *PLoS Comput Biol* 20(7): e1012274. <https://doi.org/10.1371/journal.pcbi.1012274>

Editor: Feng Fu, Dartmouth College, UNITED STATES OF AMERICA

Received: December 22, 2023

Accepted: June 24, 2024

Published: July 11, 2024

Copyright: © 2024 Guo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data and analysis code that replicate all results in the manuscript and [supplemental materials](#) are also openly shared and are available at: <https://doi.org/10.5281/zenodo.11239516>.

Funding: This work was supported by the National Natural Science Foundation of China (32271110 to ZW) https://www.nsf.gov.cn/english/site_1/index.html and the Tsinghua University Initiative Scientific Research Program (20235080047 to ZW) <https://www.tsinghua.edu.cn/en/Research.htm> Funders did not play any role in the study

Abstract

Altruistic punishment is key to establishing cooperation and maintaining social order, yet its developmental trends across cultures remain unclear. Using computational reinforcement learning models, we provided the first evidence of how social feedback dynamically influences group-biased altruistic punishment across cultures and the lifespan. Study 1 ($n = 371$) found that Chinese participants exhibited higher learning rates than Americans when socially incentivized to punish unfair allocations. Additionally, Chinese adults showed slower learning and less exploration when punishing ingroups than outgroups, a pattern absent in American counterparts, potentially reflecting a tendency towards ingroup favoritism that may contribute to reinforcing collectivist values. Study 2 ($n = 430$, aged 12–52) further showed that such ingroup favoritism develops with age. Chinese participants' learning rates for ingroup punishment decreased from adolescence into adulthood, while outgroup rates stayed constant, implying a process of cultural learning. Our findings highlight cultural and age-related variations in altruistic punishment learning, with implications for social reinforcement learning and culturally sensitive educational practices promoting fairness and altruism.

Author summary

How do people from different cultures and ages learn to punish unfair behavior? Such punishment is crucial for societal cohesion, yet remains poorly understood. We tested participants from China and the USA, and from early adolescence to adulthood. Chinese individuals, compared to Americans, quickly associate punishing unfair actions with social feedback. However, they tend to hesitate when learning to punish ingroup members, unlike American individuals. Interestingly, this preference for ingroup members becomes more prominent as Chinese participants grow from adolescence into adulthood, reflecting a gradual adoption of collectivist values. These findings emphasize how cultural norms and developmental stages strongly influence the learning processes that shape social decision-making, offering insights for creating interventions that promote justice and cooperation in diverse societies.

design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Human cooperation, particularly the tendency to collaborate anonymously with unrelated individuals in large groups, presents a fascinating evolutionary conundrum. Central to understanding this phenomenon is altruistic punishment, wherein bystanders incur costs to punish cooperative norm violators without material gain. This mechanism is essential in promoting social harmony, fairness, and order [1–4]. Notably, altruistic punishment often exhibits parochialism amidst intergroup conflicts; adults typically mete out harsher punishments for outgroup violators, and are more likely to advocate for ingroup victims [5–8]. These biases underscore the profound impact of our group-centric nature on moral standards. However, empirical evidence is lacking in showing how these biases are learned and influenced by social norms across cultures and throughout the lifespan [9], although the norm-psychology account emphasizes the role of culturally transmitted cooperative norms and cognitive mechanisms [10,11]. Investigating the developmental trajectory of altruistic punishment can yield crucial insights into the nature of human morality, norms, and strategies to foster altruistic norm enforcement.

Social norms, as shared behavioral standards, regulate the expected conduct within a community [12–14]. The norm-psychology account posits that culture-specific norms, which individuals internalize and adhere to, influence parochialism in social behaviors [15]. For instance, in individualistic-oriented cultures, where egalitarian values are socially incentivized [2,16], people typically mete out equal punishments to unfair ingroup and outgroup members when making decisions with deliberate consideration [7,17]. In contrast, collectivist-oriented cultures, such as China, place greater importance on obligations to prioritize groups' interests and maintain ingroup harmony [18–21]. Consequently, Chinese participants have been observed punishing ingroup transgressors less severely than outgroup offenders despite their negative attitudes toward unfair ingroups [9]. These findings highlight the significant influence of culturally specific norms on group bias in altruistic punishment. However, the learning processes underlying these behaviors, especially how individuals discern the appropriate response to norm violations during intergroup interactions remains an open question.

Previous research indicates that individuals learn social norms based on social feedback, including others' suggestions, evaluations, and affects [22–24], and adjust their altruistic punishment accordingly [22,25]. Meanwhile, showing preferential treatment or favoring ingroup members often receives substantial positive social feedback, such as good reputations [26] and increased collaborations [27]. In individualistic cultures, principles of fairness and equality are often more highly valued and socially rewarded [28]. Consequently, anti-discrimination behaviors across groups are reinforced. In contrast, in collectivist cultures, individuals are often educated and socialized to prioritize the needs and interests of the group (e.g., family, school, society) over their personal interests [29]. For example, adolescents who prioritize the needs and interests of their group are highly appreciated by their parents, educators, and peers [30]. Such feedback serves as a potent reinforcement mechanism, encouraging individuals to consistently favor their own group, making it a deeply ingrained and socially reinforced behavior [31–33].

Notably, according to the norm-psychology account, once individuals have internalized their culturally ingrained norms, such norms become ends in themselves or part of individuals' utility functions and motivate action regardless of other payoffs and sanctions [34], which can hinder the process of learning new social norms [31]. Research has shown that for those who have firmly internalized a social norm and developed specific behavioral tendencies, violating these norms can cause psychological discomfort, even when there are clear material benefits to doing so [35]. In addition, to maintain a positive self-image, individuals avoid situations that

may cause them to deviate from their internalized social norms and established behavioral preferences [35]. Consequently, when individuals have established a stable ingroup bias for punitive behavior, external social feedback that encourages punishment of ingroup transgressors is less likely to be learned. This effect may be more pronounced in collectivist cultures that prioritize group harmony and group interests. However, although previous research has shed light on how internalized norms may block the adoption of new social norms [31,34,35], such studies have predominantly employed condition-comparison methods in behavioral tasks, which cannot capture the temporal dynamics of learning processes. Therefore, the current study primarily aims to examine how individuals from different cultural backgrounds dynamically learn to impose altruistic punishment in response to external social feedback and how their pre-existing biases affect the learning process.

To achieve this goal, we employed a computational approach using Reinforcement Learning (RL) models. Reinforcement learning theory illustrates how decisions are paired with outcomes over time, and elucidates how learning occurs via prediction error, the discrepancy between expected and actual outcomes. Prediction error is scaled by the learning rate, which indicates the extent to which an individual updates their beliefs or expectations based on new information or feedback [36,37]. By analyzing learning rates, we can delineate quantitative differences in how individuals learn to punish unfair ingroup and outgroup members, and illustrate how these learning mechanisms vary across cultures [38–41]. Furthermore, following recent studies [36,42] we can capture the pre-existing action bias (a bias to make a response that inhibits punishment of ingroup members, regardless of its expected outcome value), and examine how it impacts the learning process. Accordingly, we hypothesized that Chinese adults might exhibit a pronounced pre-existing ingroup action bias, reflecting their internalized cultural norms of ingroup preferences. Consequently, we anticipated that they might demonstrate lower learning rates when learning to punish ingroup (vs. outgroup) members. In contrast, influenced by individualistic values that emphasize fairness, Americans might exhibit a less pronounced pre-existing ingroup action bias. Therefore, we predicted that they would likely demonstrate similar learning rates for both ingroup and outgroup transgressions.

Another unaddressed question pertains to the developmental trajectory of altruistic punishment during intergroup interactions. Previous research suggests that in individualistic-oriented cultures, group bias in altruistic punishment may decrease with age [43,44], as children learn social norms against bias and discrimination [43,45,46]. For example, American children's ingroup favoritism in altruistic punishment appears to decrease between the ages of 6 and 8 [44]. Additionally, older American children (age 9 and above) are more likely than younger children (age 6 and under) to rectify existing unequal resource allocations between socially disadvantaged outgroups (i.e., African Americans) and advantaged ingroups (i.e., European Americans), thereby promoting intergroup fairness [47]. This suggests that if social norms prioritize intergroup fairness over ingroup favoritism, group bias in altruistic punishment may diminish with age due to the socialization process. This hypothesis warrants further investigation, especially since there is currently a reliance primarily on WEIRD (Western, Educated, Industrialized, Rich, Democratic) samples [43,45,46]. The developmental trajectory of altruistic punishment may differ in collectivist-oriented cultures such as China, where ingroup loyalty is more internalized and socially incentivized than egalitarian values. Moreover, as individuals age, they are likely to become more integrated into their cultures [48], absorbing and reinforcing the norms and values that emphasize ingroup harmony and cohesion [29]. Consequently, we hypothesize that for Chinese participants, the pre-existing ingroup bias for punishing ingroup members may intensify with age, resulting in a decreased learning rate for punishing ingroup members. This phenomenon may reflect an increased reluctance to act against one's ingroup due to the greater internalization of collective values over time.

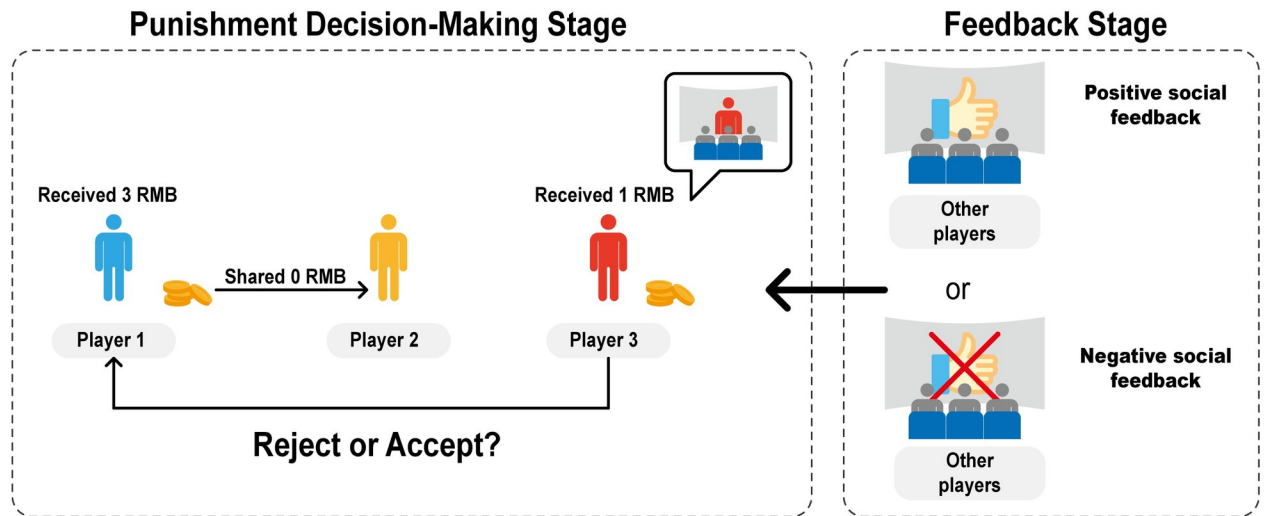


Fig 1. Reinforcement Learning Altruistic Punishment Task. Participants selected a team color (yellow) and were assigned as Player 3, with Players 1, 2, and observers computer-generated. As shown in the left panel, during the punishment decision-making stage, Player 1 chose to share 0 RMB with Player 2. Endowed with 1 RMB, Player 3 then chose to either accept or reject the allocation, with rejection equating to a costly punishment. As shown in the right panel, in the feedback stage, Player 3 received social feedback from the observers in the waiting room, conveyed by a thumbs-up for positive feedback or a thumbs-up overlaid with a red cross for negative feedback. The feedback was based on a 60% approval threshold. Participants were assigned to either the ‘punishment-encouragement’ or ‘acceptance-encouragement’ condition, incentivizing either punishment or acceptance with an 80% chance of positive feedback and a 20% chance of negative feedback, respectively.

<https://doi.org/10.1371/journal.pcbi.1012274.g001>

Furthermore, while research has highlighted cultural differences in the adoption of altruistic punishment, research gaps remain regarding the role of individual differences in explaining these learning differences. Previous work has shown that the process by which individuals learn and enforce social norms can be significantly influenced by the extent to which they perceive themselves as belonging to a particular social group, a concept referred to as group identity salience [49,50]. Empirical research indicates that individuals in collectivist cultures, which prioritize group interests, are more likely to identify with their group than those in individualistic cultures [51]. Furthermore, individuals with strong group identification are more likely to act in accordance with internalized group norms [52] and are reluctant to adopt social norms that conflict with their internalized norms [53–55]. In contrast, individuals who are willing to abandon norms are often less identified with their groups [56]. Therefore, we hypothesize that individual differences in group identity may explain the observed cultural differences in learning altruistic punishment. Specifically, compared to American participants, Chinese participants are more likely to form a strong group identity as a result of experimental manipulation involving group membership, and consequently exhibit a culturally ingrained in-group bias in punishment. Such a pre-existing bias may further reduce the learning rates of social feedback among Chinese participants relative to their American counterparts.

To examine how learning to impose altruistic punishment varies across cultures and throughout the lifespan, we devised a novel reinforcement learning altruistic punishment paradigm (Fig 1). In this framework, participants (acting as third parties) could either accept observed allocations or punish unfair players at a cost. The unfair players may be members of the participant’s own ingroup or outgroup. This distinction is created through the minimal group paradigm, which prompts participants to choose to join a team based on their color preference [50,57]. Such minimal groups are devoid of the emotional bonds, learning opportunities, and historical contexts that characterize real-world social groups, thus helping to control for extraneous variables. Despite their simplicity, research has demonstrated that these

groups effectively reveal individuals' general ingroup bias, which predicts ingroup favoritism in various real-life groups [57,58]. Subsequently, participants received trial-by-trial feedback on other observers' evaluations of their punitive decisions. In the "punishment-encouraged" condition, punishing had an 80% likelihood of receiving positive feedback (thumbs up) from other players, while accepting only had a 20% likelihood. In the "acceptance-encouraged" condition, these associations were reversed. This design allowed us to observe how individuals dynamically adjusted their behavior based on social feedback, with a particular focus on responses to unfair allocations committed by ingroup versus outgroup members. Crucially, these contingencies were not explicitly instructed but were implicitly learned through trial and error, enabling participants to update their decisions based on feedback from other observers. Using computational modeling, we analyzed the dynamic adjustments people made to altruistic punishment during intergroup interactions based on external social feedback. This novel approach stands apart from previous reinforcement learning paradigms in which participants learn arbitrary associations between actions and outcomes, such as associating a picture with rewards. In contrast, the associations in our study have real-life social significance, such as encouraging the punishment of ingroup selfishness or not, which taps into individuals' internal values regarding group identity and fairness. By integrating the behavioral task with computational modeling, we developed a tool to probe into the dynamic reinforcement learning mechanisms that underpin norm enforcement across different cultures and age groups—an area that remains relatively uncharted.

Addressing this knowledge gap, we conducted two studies. In Study 1, we compared reinforcement learning parameters between Chinese and American participants, quantifying how manipulated social feedback interacts with cultural norms to shape individuals' altruistic punishment behaviors. Study 2 further investigated the developmental trajectory of group bias in altruistic punishment among Chinese participants, spanning from adolescence to adulthood, especially considering the limited data on non-WEIRD samples. We hypothesized that: (1) learning to punish ingroup and outgroup members could be computationally differentiated; (2) during the learning process, both the pre-existing action bias and learning rates would show more pronounced group distinction among Chinese participants than among American participants; (3) ingroup bias in altruistic punishment among Chinese participants would increase with age, while the learning rate for punishing ingroups would decrease with age; and (4) The observed cultural differences in learning altruistic punishment may be explained by individual differences in group identity.

Results

The study recruited 389 adults (217 Chinese, 172 American) online and 213 adolescents offline (see [Methods](#); total $N = 602$). Study 1 investigated whether the learning mechanisms underlying ingroup bias in punishment behavior differed across cultures. We analyzed participants' altruistic punishment towards ingroup and outgroup members among the Chinese and American participants, using computational reinforcement learning models to estimate learning rates (α) and temperature (β) parameters, as well as the constant bias term. Study 2 further examined the development of ingroup bias across different age groups among Chinese participants.

Chinese participants show greater ingroup bias in altruistic punishment than American participants

First, we assessed the influence of culture and the divider's group membership on participants' punishment decisions (0 = accept, 1 = punish) during the pre-test stage using a GLMM model.

Fixed effects were culture and the divider's group membership, and by-participant random intercepts, as well as the random slope for divider's membership, were also included to further reduce the probability of committing false-positive errors (see [Method](#) for the detailed model selection procedure and [S1 Table](#) for the model selection process). Additionally, to control for potential confounding demographic differences between American and Chinese samples, we included age, gender, education level, and subjective socioeconomic status as covariates in the analyses. Results revealed a significant interaction between culture and divider groups ([Fig 2A](#); see [S2 Table](#) for the complete model results). Pairwise comparisons showed that both Chinese and American participants punished ingroup members less severely compared to outgroup members (Chinese: $b = -1.44$, $SE = 0.19$, $z = -7.69$, $p < .001$; American: $b = -0.61$, $SE = 0.22$, $z = -2.80$, $p = .026$). Notably, Chinese participants exhibited a more pronounced ingroup bias compared to their American counterparts ($b = -0.82$, $SE = 0.29$, $z = -2.94$, $p = .003$). This finding aligns with previous research [7] and indicates the role of cultural values in shaping group biases in altruistic punishment. In addition, we assessed participants' beliefs about the game's authenticity to rule out the possibility that differences in these beliefs might contribute to different levels of ingroup bias across groups. Analysis revealed no significant differences in beliefs about the game's authenticity among culture groups ($F(2, 599) = 0.27$, $p = .765$, $\eta_p^2 = 0.001$; Chinese adults: $M = 3.91$, $SD = 1.70$; American adults: $M = 3.87$, $SD = 1.61$; Chinese adolescents in study 2: $M = 3.78$, $SD = 1.98$).

Next, we investigated whether participants adjusted punitive behavior in response to social feedback during the reinforcement learning stage, and whether this learning was contingent on trial-by-trial feedback. Average punishment rates were computed across trials, as a function of culture (Chinese vs. American), divider group membership (ingroups vs. outgroups), and social feedback (punishment-encouragement vs. acceptance-encouragement) ([Fig 2B and 2C](#)). Participants from both cultures displayed higher punishment rates than chance level (50%; all $t_s > 2.28$, $p_s < .05$) when rewarded to punish and lower punishment rates than chance level when rewarded to accept (50%; all $t_s < -8.28$, $p_s < .001$). Further analysis of the trial-by-trial punishment rates showed a positive association with the trial number when punishment was rewarded ($b = 0.04$, $SE = 0.01$, $t = 3.11$, $p = .003$) and a negative association when acceptance was rewarded ($b = -0.03$, $SE = 0.01$, $t = -3.90$, $p < .001$) ([Fig 2D](#)). Taken together, these results suggest that individuals can assimilate punishment norms through social feedback.

Then we examined potential differences in punishment behavior across cultures, divider groups, and punishment norms during the reinforcement learning stage (see [S3 Table](#) for the model selection process and [S4 Table](#) for the complete model results). A GLMM analysis uncovered a significant interaction between cultures and divider groups ($b = -0.49$, $SE = 0.24$, $z = -2.03$, $p = .042$) ([Fig 2B and 2C](#)). Pairwise comparisons showed that Chinese participants punished ingroup dividers less severely compared to outgroup dividers, regardless of whether punishment ($b = -0.87$, $SE = 0.22$, $z = -4.01$, $p < .001$) or acceptance was rewarded ($b = -0.91$, $SE = 0.23$, $z = -4.06$, $p < .001$). In contrast, no significant differences were found between punishment towards ingroup and outgroup dividers among American participants, whether punishment was rewarded ($b = -0.61$, $SE = 0.26$, $z = -2.37$, $p = .084$) or acceptance was rewarded ($b = -0.20$, $SE = 0.26$, $z = -0.76$, $p = .871$). Consistent with the pre-test stage, the ingroup bias in punishment behavior was more pronounced among Chinese participants during the reinforcement learning stage. These results highlight that despite being exposed to a similar new environment and receiving identical social feedback, individuals from different cultural backgrounds may interpret or respond differently due to their ingrained cultural norms, which persistently influence their learning processes.

Nevertheless, solely focusing on mean punishment behavior does not elucidate how ingroup bias unfolds throughout the learning process. On the one hand, regarding the social

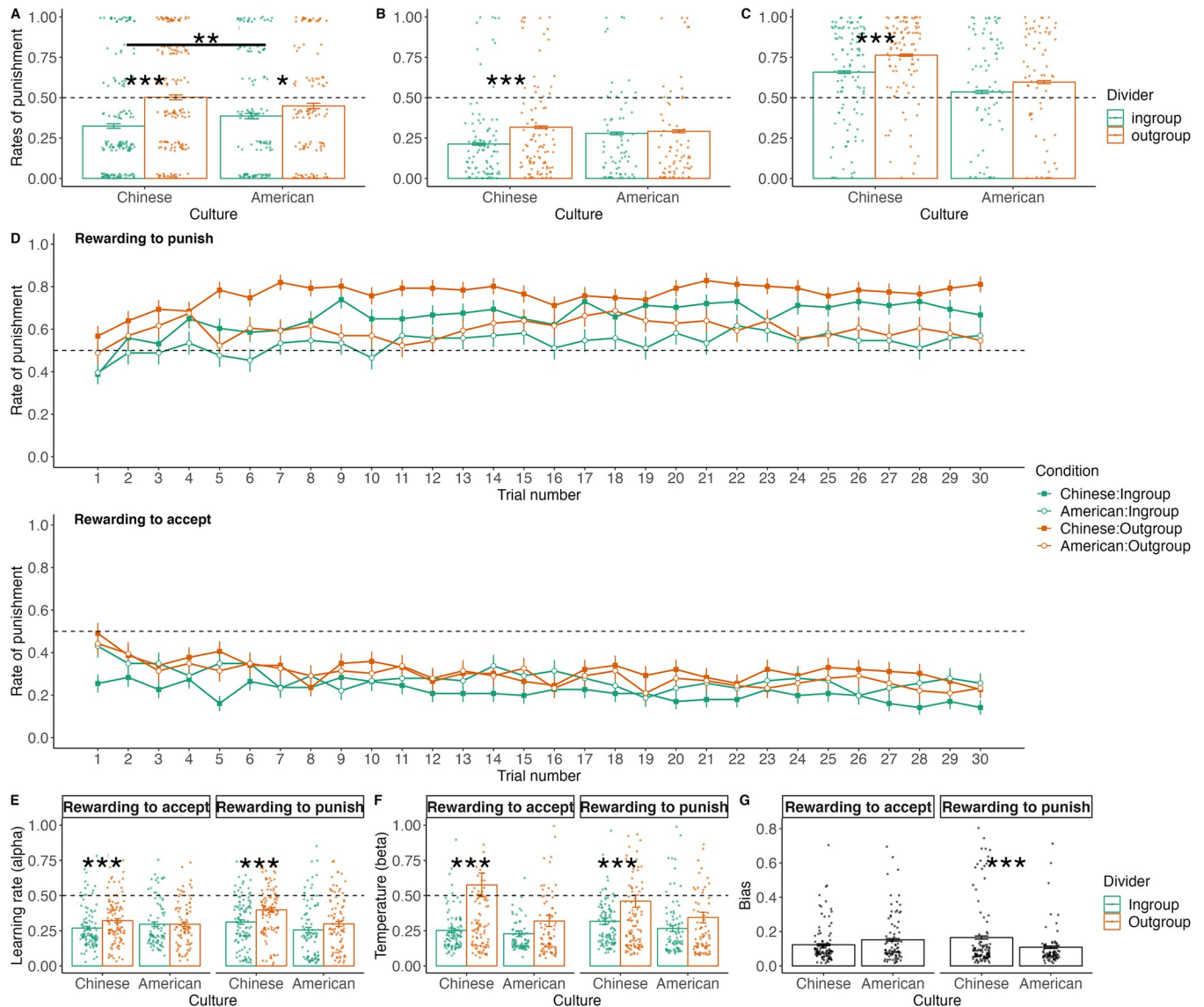


Fig 2. Cultural and group influences on altruistic punishment behavior in pre-test and reinforcement learning stages. **A** Pre-test stage: Chinese participant exhibited stronger ingroup bias compared to American participants. Overall, participants punished ingroup (green) more severely compared to outgroup (orange) dividers. However, Chinese participants exhibited an enhanced group bias ($b = -0.82, SE = 0.29, z = -2.94, p = .003$). **B** Reinforcement learning stage: when acceptance was rewarded, Chinese participants punished ingroup dividers (green) less severely compared to outgroup dividers ($b = -0.91, SE = 0.23, z = -4.06, p < .001$), no significant differences were found between punishment towards ingroup and outgroup dividers among American participants ($b = -0.20, SE = 0.26, z = -0.76, p = .871$). **C** Reinforcement learning stage: when punishment was rewarded, Chinese participants punished ingroup dividers (green) less severely compared to outgroup dividers ($b = -0.87, SE = 0.22, z = -4.01, p < .001$), no significant differences were found between punishment towards ingroup and outgroup dividers among American participants ($b = -0.61, SE = 0.26, z = -2.37, p = .084$). **D** Group level trial-by-trial punishment rates in the two divider conditions (ingroup: green, outgroup: orange) for each culture group (Chinese: squares, American: circles). Points represent group mean; error bars are standard errors. **E** Regardless of whether punishment or acceptance was rewarded, Chinese participants had lower learning rates for ingroup norms ($b = -0.07, SE = 0.01, t = -5.51, p < .001$), while American participants showed no differences ($b = -0.02, SE = 0.01, t = -1.46, p = .460$). **F** Regardless of whether punishment or acceptance was rewarded, Chinese participants had lower β when learning ingroup norms ($b = -0.23, SE = 0.04, t = -5.85, p < .001$), while no differences were observed for American participants ($b = -0.08, SE = 0.04, t = -1.90, p = .231$). **G** Chinese participants showed a stronger bias compared to American participants, especially when punishment was rewarded ($M_{Diff} = 0.06, SE = .01, t = 5.11, p < .001$). Error bars represent standard errors. * $p < .05$, *** $p < .001$.

<https://doi.org/10.1371/journal.pcbi.1012274.g002>

feedback learning process, the observed fewer punishment behaviors towards ingroups could arise from distinct interpretations of external feedback. This could manifest as slower updating of values to punish ingroups (indicated by lower learning rates α). Alternatively, despite

learning values from the environment, participants may follow the learned values less closely (indicated by higher temperature β). On the other hand, beyond the social feedback learning process, individuals may demonstrate a stable action bias that leads to leniency towards ingroups, showing a resistance to learning social feedback. To dissect these possibilities, we conducted computational modeling analyses to inspect the underlying mechanisms.

Learning to punish ingroup and outgroup selfishness differs computationally

We fitted computational models of reinforcement learning to estimate the parameters of learning rates (α) and temperature (β), as well as the constant bias term (see [Methods](#)). The learning rate (α) captured how individuals learned and modified their value expectations concerning punishments based on external feedback. Temperature (β) reflected how individuals adjusted their behavior during decision-making based on these value expectations. Furthermore, to depict individuals' deeply ingrained cultural preferences, which were resistant to changes in punishment based on social feedback, we integrated a persistent bias term into our model. This bias term was designed to capture the default bias of individuals that showed resistance to social feedback. Then we employed the Integrated Bayesian Information Criterion (Integrated BIC) to compare seven candidate models (see [Methods](#)). In samples combining both Chinese and American participants, as well as in separate analyses, a model with distinct α for ingroup and outgroup dividers across two blocks, separate β for dividers, and a bias term ($4\alpha 2\beta + \text{bias model}$) provided the best fit for participants' choices (see [S5 Table](#)), suggesting a computationally distinct learning process for punishing ingroups versus outgroups. To further validate our computational model, we conducted a parameter recovery (see [Methods](#)) using the winning $4\alpha 2\beta + \text{bias model}$, thereby demonstrating the recoverability of the parameters (see [Fig 3A](#) and [S5 Table](#)).

Moreover, although a higher learning rate typically suggests that learning is driven mainly by recent feedback, and a higher temperature implies more randomness across trials, emerging evidence increasingly suggests that the relationship between these parameters and behavioral choices is heavily dependent on the experimental design [59–61]. This suggests that the extent to which higher learning rates led to increased punishment in contexts where social feedback encouraged such responses or, conversely, to decreased punishment in contexts that encouraged acceptance, remained an open question. Therefore, building on previous research, we utilized simulated experiments and compared empirical data from our experiments to enable a clear interpretation of the relationship between parameters (learning rate and temperature) and punishment behaviors across different experimental conditions. By simulating data from 5000 participants using the $4\alpha 2\beta$ model with a bias term under both conditions of punishment-encouragement and acceptance-encouragement, we produced 10,000 simulated participants and 40,000 α values, fully encompassing the potential range from 0 to 1. We calculated the punishment rates as the percentage of punishment enforced by the same participant, averaged across blocks for each divider group.

Our findings revealed that simulated α values exerted quadratic impacts on punishment rates in conditions that encouraged punishment. Specifically, in the punishment-encouragement condition, the simulation experiment demonstrated a trend where punishment rates first increased and then decreased as a function of the learning rate (α), with an optimal learning rate of 0.64 maximizing punishment rates (see [Fig 3B](#)). However, when further examining the relationship between learning rates and rates of punishment in the empirical data, we found the average α values among Chinese ($M = 0.28$, $SD = 0.24$) and American ($M = 0.21$, $SD = 0.24$) adults in study 1 and Chinese adolescents in study 2 ($M = 0.41$, $SD = 0.23$) were

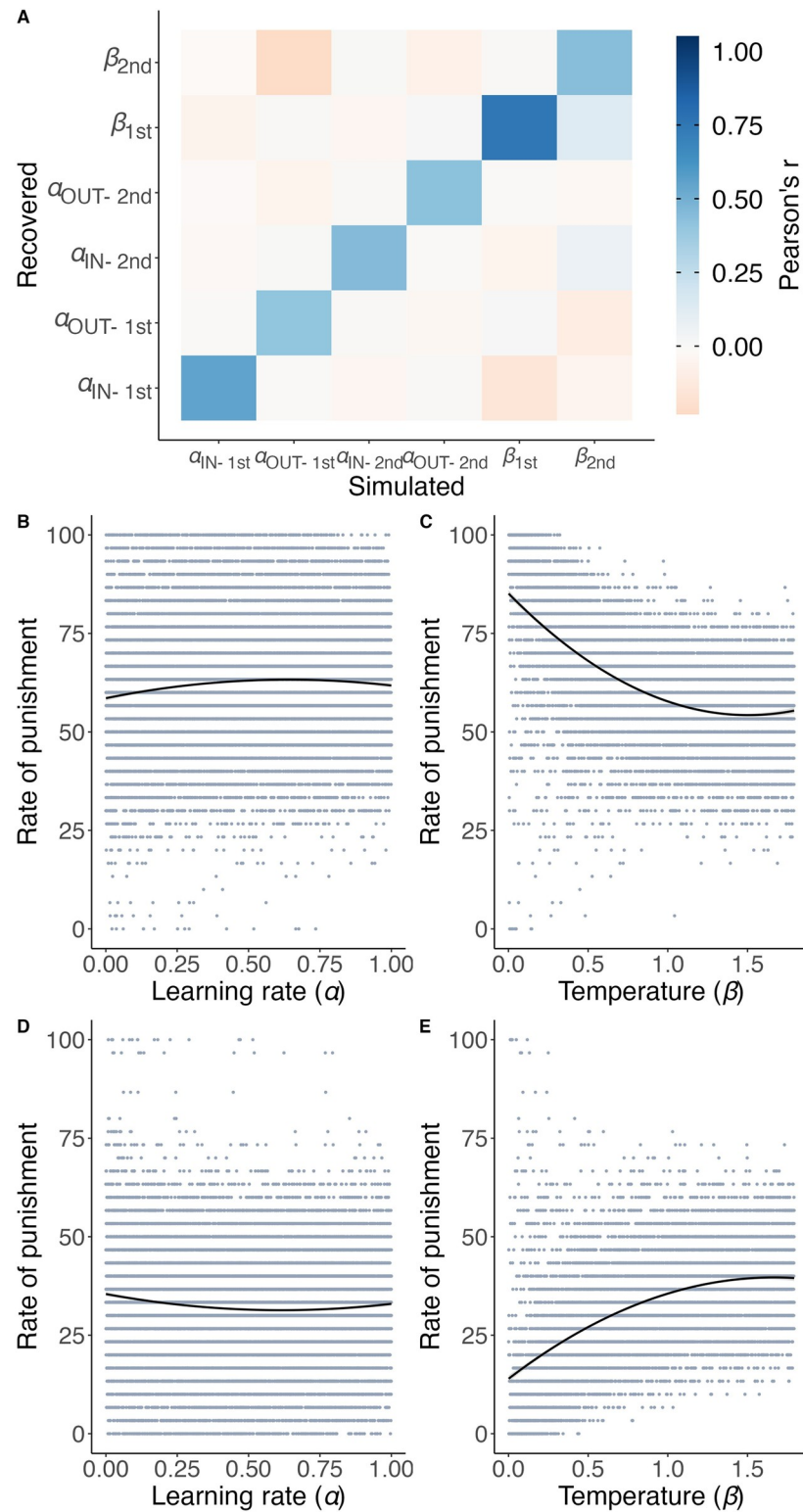


Fig 3. Computational modeling reveals distinct norm learning and decision-making processes in punishing ingroups and outgroups. **A** Parameter recovery analysis for the winning $4a2\beta + \text{bias}$ model demonstrates the recoverability of the parameters. Data from 15,625 simulated participants were used, with a confusion matrix depicting correlations between simulated and fitted parameters. Enhanced colors represent higher values. See [S6 Table](#) for the source data. **B** When punishment was rewarded, a simulation experiment illustrated the quadratic relationship

between α and the rate of punishment. An α of 0.64 was identified as the point at which the rate of punishment was maximized. C Quadratic relationship between β and the rate of punishment. A β of 1.51 was identified as the point at which the rate of punishment was minimized. D When acceptance was rewarded, a simulation experiment illustrated the quadratic relationship between α and the rate of punishment. An α of 0.61 was identified as the point at which the rate of punishment was minimized. E Quadratic relationship between β and punishment rates. A β of 1.66 was identified as the point at which the rate of punishment was maximized.

<https://doi.org/10.1371/journal.pcbi.1012274.g003>

below the optimal learning rate of 0.64. In fact, 85.3% of participants' α values in study 1 and 79.4% of participants' α values in study 2 were found to be below 0.64. These results demonstrated that in our empirical study, a higher learning rate represented faster adaptation to increased punishment in response to external feedback encouraging such behavior. These results were further supported by the positive correlation between higher α values and punishment rates in empirical data for adults in study 1 ($r = 0.29, p < .001$) and adolescents in study 2 ($r = 0.26, p < .001$).

Conversely, in conditions that encouraged acceptance, the simulated experiment revealed an initially declining and then increasing trend in punishment rates with increasing α , indicating a learning rate of 0.61 that minimized punishment rates (see Fig 3D). Meanwhile, analysis of empirical data elucidated that the mean α values for both Chinese ($M = 0.30, SD = 0.19$) and American ($M = 0.30, SD = 0.21$) adults in study 1, and Chinese adolescents in study 2 ($M = 0.42, SD = 0.23$), were below the learning rate of 0.61. Notably, 88.7% of participants' α values in study 1 and 81.0% of participants' α values in study 2 were below this threshold. Moreover, the higher α values were negatively related to punishment rates under conditions encouraging acceptance (study 1: $r = -0.21, p < .001$; study 2: $r = -0.14, p < .001$), suggesting that a higher learning rate might reflect quicker adjustment to reduce punishment according to external feedback.

Next, when examining the relationship between temperature and punishment rates, we found quadratic effects of temperature on punishment rates (see Fig 3C). Under conditions that encouraged punishment, simulations showed punishment rates decreasing and then increasing with rising β , indicating a β of 1.51 for minimizing punishment. Empirical data revealed mean β values for both Chinese ($M = 0.39, SD = 0.37$) and American ($M = 0.31, SD = 0.28$) adults in study 1, and Chinese adolescents in study 2 ($M = 0.42, SD = 0.23$) below the 1.51 threshold, with 98.2% of the β values in study 1 and 93.8% of the β values in study 2 under this threshold. Temperature was negatively correlated with punishment rates (study 1: $r = -0.08, p < .001$; study 2: $r = -0.27, p < .001$), suggesting higher temperatures decrease the likelihood of actions aligning with social feedback that encourages punishment.

Finally, under conditions that encouraged acceptance, the simulation showed punishment rates rising then falling with temperature (β), peaking at a β of 1.66 (see Fig 3E). Empirical findings indicated average β values for Chinese ($M = 0.30, SD = 0.19$) and American ($M = 0.30, SD = 0.21$) adults in study 1, and Chinese adolescents in study 2 ($M = 0.42, SD = 0.23$) below the 1.66 threshold, with 97.4% of the β values in study 1 and 93.2% of the β values in study 2 under this threshold. There was a significant positive relationship between temperature and punishment rates (study 1: $r = 0.23, p < .001$; study 2: $r = 0.46, p < .001$), suggesting that higher temperatures reduced the likelihood of decisions consistent with social feedback that encourages acceptance.

Taken together, these results revealed that in our research, higher learning rates represented accelerated social feedback learning, allowing quicker adaptation in punishment or acceptance as encouraged. Higher temperatures were correlated with increased behavioral randomness, implying a rise in decision-making variability.

Chinese participants show lower learning rates for punishing ingroups than American participants

We examined the impact of culture, divider groups, and punishment norms on learning rates (α) and temperature (β) in reinforcement learning. Regarding α , faster learning occurred in the first block ($M = 0.36$, $SD = 0.24$) than in the second ($M = 0.26$, $SD = 0.19$; $t = -8.75$, $p < .001$). Linear mixed model (LMM) analyses revealed a significant two-way interaction between divider group membership and cultural background across two blocks ($b = -0.09$, $SE = 0.04$, $t = -2.58$, $p = .010$) (Fig 2E; see S7 Table for the model selection process and S8 Table for the complete model results). Chinese participants had lower learning rates for ingroup norms ($b = -0.07$, $SE = 0.01$, $t = -5.51$, $p < .001$), while American participants showed no significant differences ($b = -0.02$, $SE = 0.01$, $t = -1.46$, $p = .460$).

Regarding β , we also found a significant two-way interaction between divider group membership and cultural background ($b = -0.14$, $SE = 0.06$, $t = -2.48$, $p = .014$) (Fig 2F; see S9 Table for the model selection process and S10 Table for the complete model results). Chinese participants had lower β for ingroup norms ($b = -0.23$, $SE = 0.04$, $t = -5.85$, $p < .001$), while no differences were observed for American participants ($b = -0.08$, $SE = 0.04$, $t = -1.90$, $p = .231$). These results uncover distinct learning mechanisms underpinning punitive behavior between Chinese and American participants, despite both showing an ingroup bias in their behavioral choices. When learning how to punish ingroup (vs. outgroup) members, Chinese participants adjusted their punishment slower in response to external feedback, as evidenced by their lower learning rates. They also demonstrated a less random decision-making process with a tendency to consistently select options with the highest expected value for ingroup members, reflected in their lower β values. This value-driven strategy in decision-making involving ingroup members among Chinese adults contributes to the observed ingroup bias in punishment behavior.

Furthermore, we examined the impact of culture and punishment norms on the bias term, which shifts individuals' preference towards accepting ingroup members' unfair allocations and thus captures individuals' default bias that exhibits resistance to social feedback. We found a significant interaction between culture and punishment norms ($F(1, 385) = 7.25$, $p = .007$, $\eta_p^2 = 0.018$; see Fig 2G). Under the punishment-encouragement condition, Chinese participants showed a stronger bias compared to American participants ($M_{\text{Diff}} = 0.06$, $SE = .01$, $t = 5.11$, $p < .001$), whereas the bias was not significantly different across cultures under the acceptance-encouragement condition ($M_{\text{Diff}} = -0.03$, $SE = .01$, $t = -2.55$, $p = .053$). These results indicated that Chinese (vs. American) participants exhibited more resistance to social feedback when they were rewarded for punishing ingroups.

Taken together, our data suggested that, compared to their American counterparts, Chinese adults demonstrated a consistent preference in treating ingroup members and were less inclined to adjust their decisions in response to dynamic external feedback about ingroup members. These findings revealed the potential influence of ingrained cultural norms on the blocking of learning from social feedback. Given that theoretical and empirical research indicates cultural norms are gradually internalized and acquired throughout development, it is plausible that such ingroup bias in learning is nurtured alongside individual development within a specific culture. Specifically, if Chinese participants increasingly internalize collectivist cultural norms emphasizing ingroup preference, this results in a more deeply rooted ingroup bias. Consequently, their responsiveness to external feedback regarding the punishment of ingroup members decreases with age, leading to a decrease in learning rates for such punitive measures. On the contrary, due to a lack of cultural norms on how to treat outgroup members, we expect no significant developmental changes in adjusting punitive behaviors

based on feedback, manifesting in stable learning rates for punishing outgroup members across different ages.

Chinese participants' ingroup bias in punishment increases with age from adolescence to adulthood

Study 2 examined age-related differences in the reinforcement learning processes of punitive decisions in Chinese individuals aged 12–52. First, we examined the influence of the dividers' group and participants' age on punitive decisions (0 = accept, 1 = punish) during the pre-test stage. A GLMM model revealed a significant interaction between age and the group of dividers ($b = -0.03$, $SE = 0.01$, $z = -2.62$, $p = .009$) (Fig 4A; see S11 Table for the model selection process and S12 Table for the complete model results). To unpack the two-way interaction effect, pairwise comparisons indicated a decrease in punishing unfair ingroup dividers as age increased ($b = -0.03$, $SE = 0.02$, $z = -2.05$, $p = .040$). However, age was not significantly related to punishment rates for unfair outgroup dividers ($b = -0.01$, $SE = 0.02$, $z = -0.34$, $p = .734$). These results suggest that the Chinese participants' group bias in altruistic punishment increases with age, and this increasing bias is primarily driven by ingroup favoritism rather than outgroup harshness. This pattern suggests that collectivist cultural values emphasizing ingroup favoritism become more embedded as individuals age.

We then tested age differences in punitive behavior during the reinforcement learning stage. Using a GLMM analysis, we identified a significant interaction between age and the group of dividers ($b = 0.03$, $SE = 0.01$, $z = 2.05$, $p = .041$) (Fig 4B and 4C; see S13 Table for the model selection process and S14 Table for the complete model results). Younger participants punished ingroups less than their older counterparts when acceptance was rewarded ($b = 0.03$, $SE = 0.02$, $z = 2.32$, $p = .020$), but tended to punish more when punishment was rewarded ($b = -0.03$, $SE = 0.02$, $z = -1.51$, $p = .131$). This implies that younger Chinese individuals were quicker to adapt to environmental changes in response to ingroup transgressions, and such adaptation tended to decline with age. By contrast, outgroup punishment remained stable in both conditions (punishment-encouragement: $b = -0.17$, $SE = 0.13$, $z = -1.41$, $p = .158$; acceptance-encouragement: $b = -0.02$, $SE = 0.10$, $z = -0.25$, $p = .800$). Thus, our study provides the first evidence that the prior observed ingroup bias in the literature [26,51,62] is mainly driven by increasing tolerance for ingroup norm violations with age, rather than intensifying harshness towards outgroups.

Learning rates decrease with age for punishing ingroups, but preserve for punishing outgroups

Next, consistent with Study 1, in samples combining both Chinese adults and adolescents, as well as in separate analyses, a model with distinct α for ingroup and outgroup dividers across two blocks, separate β for dividers and a bias term ($4\alpha 2\beta + \text{bias}$ model) provided the best fit for participants' choices (see S15 Table), indicating a distinct computational process for punishing ingroups and outgroups. Upon comparing the learning rates across two blocks, the first block demonstrated a significantly higher learning rate ($t = 17.34$, $p < .001$), denoting faster learning. As in study 1, we then used LMM to analyze the effects of age, group membership of dividers, and punishment norms on learning rate. LMMs revealed a significant two-way interaction between the group membership of dividers and age across two blocks ($b = -0.06$, $SE = 0.02$, $t = -2.99$, $p = .003$) (Fig 4D; see S16 Table for the model selection process and S17 Table for the complete model results). Crucially, the learning rate for ingroup punishment norms decreased with age ($b = -0.14$, $SE = 0.03$, $t = -4.21$, $p < .001$), while the α for outgroup punishment norms remained stable ($b = -0.02$, $SE = 0.03$, $t = -0.68$, $p = .500$). These results indicate that

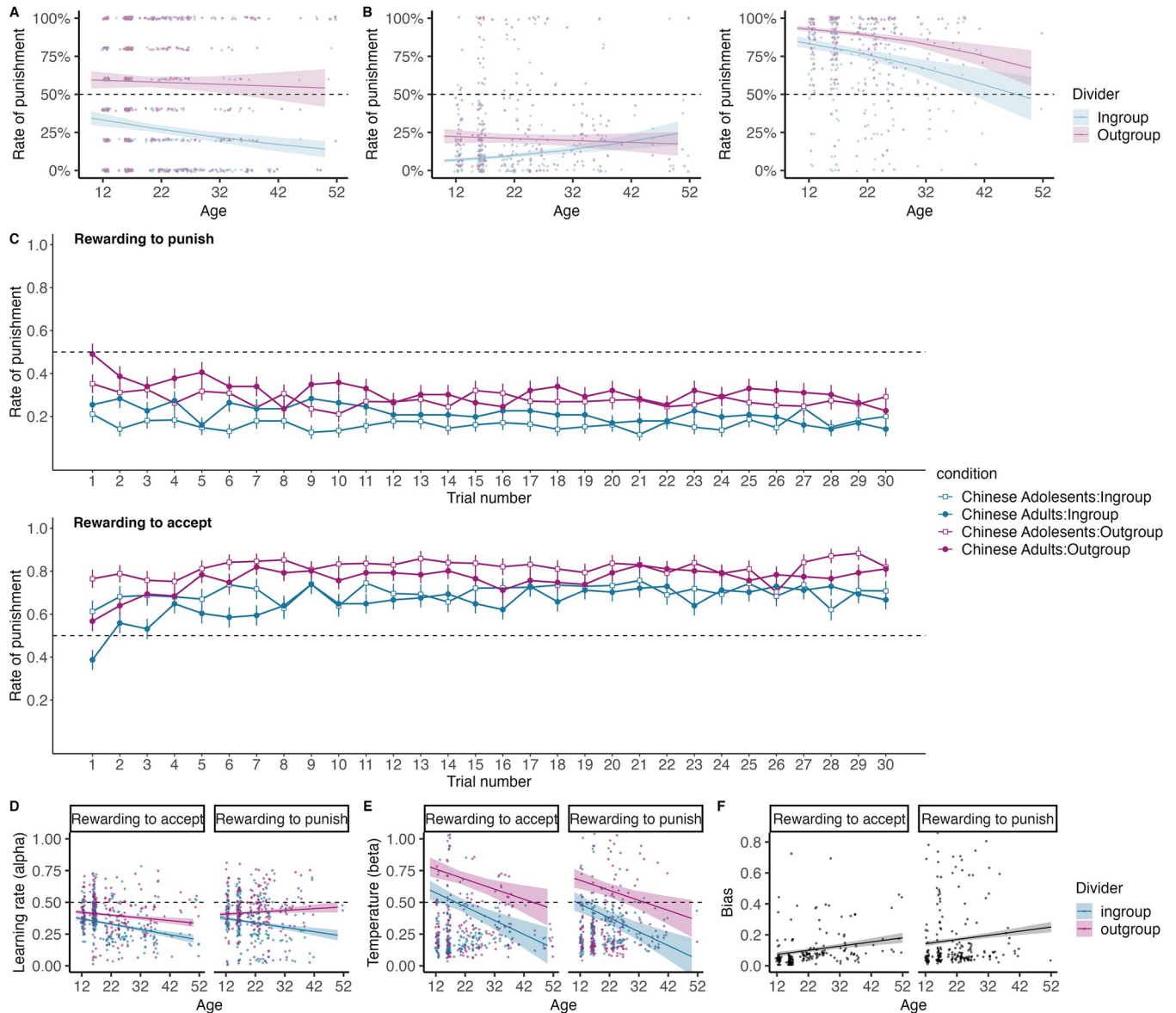


Fig 4. Age-related differences in ingroup bias and learning mechanisms among Chinese individuals. **A** Pre-test stage: the interaction effect between age and dividers' group reveals a decrease in punishing unfair ingroup dividers (blue) with increasing age ($b = -0.03$, $SE = 0.02$, $z = -2.05$, $p = .040$), while punishment rates for outgroup dividers (purple) remain unchanged ($b = -0.01$, $SE = 0.02$, $z = -0.34$, $p = .734$). **B** Reinforcement learning stage: younger participants' behavioral responses toward ingroups were more consistent with the norms than were older participants: when acceptance was rewarded, younger participants exhibited less punishment compared to older individuals ($b = 0.03$, $SE = 0.02$, $z = 2.32$, $p = .020$); when punishment was rewarded, younger participants exhibited more punishment than older participants $b = -0.03$, $SE = 0.02$, $z = -1.51$, $p = .131$). By contrast, the age differences were not significant for unfair outgroup dividers under either social feedback condition. **C** Trial-by-trial punishment rates for ingroup (blue) and outgroup (purple) dividers across age groups (adolescents: squares, adults: circles). Points represent group mean, and error bars are standard errors. **D** learning rate (α) for punishing unfair ingroups decreased with age ($b = -0.14$, $SE = 0.03$, $t = -4.21$, $p < .001$), while the α for punishing outgroups was not associated with age ($b = -0.02$, $SE = 0.03$, $t = -0.68$, $p = .500$). **E** Temperature (β): Lower β for learning ingroup norms than outgroup norms regardless of age ($b = -0.25$, $SE = 0.05$, $t = -5.36$, $p < .001$); β decreased with age ($b = -0.12$, $SE = 0.04$, $t = -2.74$, $p = .006$). **F** the bias term significantly increased with age regardless of punishment norms ($b = 0.13$, $SE = 0.05$, $t = 2.65$, $p = .008$). Error bars represent standard errors. * $p < .05$. ** $p < .01$. *** $p < .001$.

<https://doi.org/10.1371/journal.pcbi.1012274.g004>

participants learned how to punish unfair ingroups more slowly compared to outgroups, and this difference became more prominent in older participants, implying a slower adaptation in punitive behavior towards ingroup members among older participants.

Additionally, when further looking into the adolescent group (aged 12–18), results revealed adolescents' heightened sensitivity to social feedback during norm learning (see [S18 Table](#) for the model selection process and [S19 Table](#) for the complete model results). The adolescent group (aged 12–18) exhibited higher learning rates when socially incentivized to accept unfairness ($b = 0.50$, $SE = 0.11$, $t = 4.66$, $p < .001$) compared to adults (aged above 18). This enhanced learning rate was consistent across the 12–18 age group ($b = 0.01$, $SE = 0.01$, $t = 0.64$, $p = .522$). Notably, when comparing ingroup and outgroup transgressions, there was no significant difference in these learning rates for the adolescent group ($b = 0.02$, $SE = 0.02$, $z = 1.01$, $p = .312$), a contrast to adult participants. Considering that adolescence is characterized by pronounced neurodevelopment and concerns about social evaluation, this age difference may imply a sensitive window for social norm learning during adolescence.

We then investigated the influences of age, group membership, and punishment norms on temperature (β). An LMM analysis revealed only the main effect of age ($b = -0.12$, $SE = 0.04$, $t = -2.74$, $p = .006$) and dividers' group membership ($b = -0.25$, $SE = 0.05$, $t = -5.36$, $p < .001$) was significant, showing that Chinese participants' temperature was lower when learning norms regarding ingroups, and the temperature decreased with age ([Fig 4E](#); see [S20 Table](#) for the model selection process and [S21 Table](#) for the complete model results). This finding suggests that in a collectivist culture, a value-driven behavioral adjustment towards ingroup members was well-established by adolescence.

Furthermore, we examined the impact of age and punishment norms on the bias term. We found the bias significantly increased with age regardless of punishment norms ($b = 0.13$, $SE = 0.05$, $t = 2.65$, $p = .008$; see [Fig 4F](#)). These results suggested a cultural inclination towards leniency for in-group members that becomes more pronounced over time.

Taken together, Study 2 provides initial evidence suggesting the role of collectivist culture in shaping ingroup bias in punitive decisions and learning mechanisms among Chinese individuals. Prior studies have indicated that group bias in third-party punishment may be driven by both a preference for ingroup members and an exclusion of outgroup members [8,51,62]. However, it remains unclear whether, in collectivist cultures as opposed to individualistic ones, the greater group bias in third-party punishment is driven more by one aspect over the other. Our findings indicate that with increasing age, Chinese participants exhibit an enhancement of pre-existing ingroup bias, accompanied by a reduction in the learning rate for punishing ingroup members. This observation suggests an evolving internalization of collectivist cultural norms that favor ingroup preference as individuals age. Conversely, the learning rate for punishing outgroup members does not follow a similar age-related pattern. These findings may suggest that a cultural bias against outgroups is not as strongly emphasized within collectivist cultures, potentially enabling a more adaptable punitive approach towards outgroups. These insights advance our understanding of how cultural values shape social decision-making and learning processes across various developmental stages.

Group identity modulates learning from altruistic punishment

Subsequently, we tested whether individual differences in self-reported group identity could explain cultural differences of ingroup bias in learning rates. A standard mediation model was implemented using the mediation package in R. As depicted in [Fig 5](#), our findings revealed that group identity significantly predicted ingroup bias in learning rate (α -out minus α -in, $b = 0.08$, $SE = 0.04$, $t = 2.05$, $p = .041$). Furthermore, a significant indirect effect of group identity (standardized indirect effect = 0.07, 95% CI [0.003, 0.140], $p = .039$, simulation = 5,000) indicated that group identity mediated the relationship between culture and group bias in learning rates.

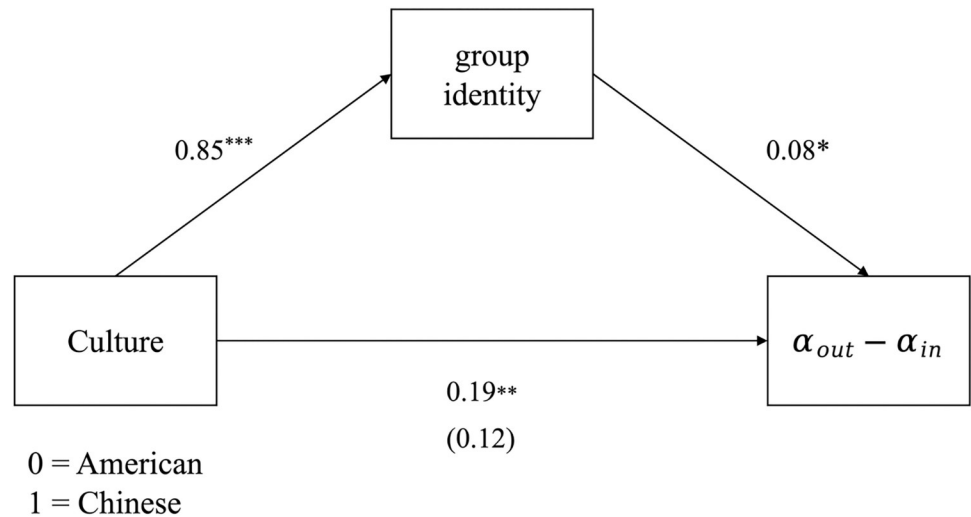


Fig 5. Indirect effect of culture on group bias in learning (α -out minus α -in) via group identity. Participants' self-reported group identity mediated the relationship between culture and group bias in learning rates (standardized indirect effect = 0.07, 95% CI [0.003, 0.140], $p = .039$).

<https://doi.org/10.1371/journal.pcbi.1012274.g005>

Discussion

Altruistic punishment is a crucial altruistic behavior that promotes cooperation. While people frequently demonstrate an ingroup bias, punishing outgroups more than ingroups when enforcing cooperative norms, how individuals dynamically adjust these biases in response to social feedback remains largely underinvestigated. In this study, we integrate behavioral tasks and reinforcement learning computational models to investigate the influence of social feedback on group-biased altruistic punishment across cultures and age groups. Our method offers a nuanced, quantitative analysis of cultural differences in learning how to punish ingroup and outgroup norm violators, and deepens our understanding of the development of group-biased altruistic punishment.

Our investigation revealed that Chinese individuals strongly adhere to ingroup loyalty norms when engaging in altruistic punishment. Analyzing pre-test punishment data without external social feedback revealed that Chinese participants showed a more significant ingroup bias in their punishment decisions compared to Western counterparts. This ingroup bias in punitive decisions among Chinese participants intensified with age, signifying a deepening process of social norm internalization over the course of individual development. These findings resonate with extant literature emphasizing the profound influence of social norms on punitive actions [22,63,64]. Within collectivist societies, individuals often receive positive social feedback, such as praise or reputation boosts, for fulfilling social roles that maintain group harmony [28]. This feedback serves as a potent reinforcement mechanism, encouraging internalized ingroup norms. As shown in prior research [7], Chinese adults reported guilt for ingroup members' violations but still exhibited less severe punitive actions upon reflection. In contrast, the cultural norms in individualistic societies like the United States, which encourage a more personal approach to morality and behavior, could lessen the impact of ingroup loyalty on altruistic punishment.

More importantly, in addition to replicating the documented ingroup bias in altruistic punitive behaviors in a non-WEIRD sample, we provided the first evidence that punishing ingroups and outgroups involves distinct *learning* mechanisms across cultures. Specifically, Chinese participants exhibited a greater pre-existing action bias that inhibits them from

punishing ingroup violators, and a group-biased learning process with lower learning rates and temperature for punishing ingroups than outgroups. In contrast, such group differentiation was not evident among American adults. These findings suggest that compared to American individuals, Chinese individuals adhered to a relatively steadfast set of ingroup loyalty norms in altruistic punishment, and such an ingrained norm impacted the learning of external social feedback regarding how to enforce punishment.

This aligns with the cultural evolutionary theory, which emphasizes the significant role of norm internalization in shaping human behavior [12,31,34,65]. The theory posits that natural selection has shaped humans to internalize norms, as such internalization decreases the costs related to information collection, processing, and decision-making, which are essential for ensuring cooperation [31,34]. However, this process of internalization also reduces behavioral flexibility, making it challenging for individuals to alter deeply ingrained norms [12,34,66]. For example, individuals frequently overestimate the difficulty of transitioning to a novel norm, with shifts occurring smoothly only in environments that provide effective feedback [35]. Additionally, to maintain a positive self-view, people typically avoid situations that might lead them away from their deeply held social norms and behaviors [35]. Although prior studies have highlighted how internalized norms might hinder adopting new norms, little is known about the temporally dynamic nature of learning processes. In contrast, our reinforcement model, which contains separate learning rates for different groups and a persistent bias term, quantified how individuals from different cultural backgrounds dynamically learn to impose altruistic punishment in response to external social feedback. It also identified how their pre-existing biases impact the learning process. Moreover, our study further supports cultural evolutionary theory by demonstrating that participants in collectivist cultures are reluctant to alter their behavior or adopt new norms when faced with social feedback that challenges the established way of treating ingroup members. It is plausible that such reluctance acts as a cultural filter, allowing only practices that align with existing values to persist and ensuring cultural stability [11,67].

Our analysis further revealed that self-reported group identity mediated this cultural difference, offering an individual-level perspective on cross-cultural variations in learning mechanisms. This finding is consistent with prior research, which indicates that the social norm learning process is markedly influenced by group identity. Specifically, Chinese participants, compared to Americans, showed a greater inclination to form group identities even within arbitrary minimal groups [7,68]. Moreover, those with a stronger identity were less inclined to adopt external social feedback regarding the punishment of ingroup (vs. outgroup) members, with heightened pre-existing ingroup bias and slower learning rates on ingroup (vs. outgroup) members. This is consistent with studies indicating that individuals with a strong identification with their groups are more likely to internalize norms [53–55] and may resist learning norms perceived as conflicting with their group's interests [54,55]. In conclusion, our study contributes to the empirical evidence of ingroup bias by revealing a computational mechanism driving group-biased norm enforcement in different cultures. This offers valuable insights for understanding and enforcing social norms cross-culturally.

Furthermore, we observed an increased ingroup bias among Chinese individuals as they aged. This was characterized by a decrease in the learning rates for punishing ingroups, which spanned from adolescence to older adulthood. Conversely, the learning rates for punishing outgroups remained stable. This finding contrasts with previous results obtained from Western cultures that suggest a reduced ingroup favoritism with age [44,45]. Our finding suggests that the collectivist culture, compared to the individualistic culture, fosters relatively more consistent internal values for ingroup norms across the lifespan. Consistent with cultural evolution theory [48,66], this trajectory suggests that individuals increasingly adopt their culture's

established norms and values with age. In the Chinese context—characterized by collectivism [68–71]—this trend indicates that older individuals, after prolonged exposure to social feedback, internalize norms favoring ingroup members more deeply. This likely reflects a cumulative social learning trajectory, where norms, reinforced through repeated interactions and feedback, may become stabilized with age [72,73]. Consequently, older adults are less likely to adjust punitive behaviors toward ingroup members in response to temporary social feedback. In contrast, the stable learning rates for punishing outgroups across age groups suggest that ingroup bias is driven more by growing favoritism toward ingroups with age rather than increasing hostility toward outgroups. This helps to clarify the theoretical debate regarding the formation of group bias [50].

Our findings also highlight adolescents' heightened sensitivity to social feedback. Given that adolescence is characterized by pronounced neurodevelopment and concerns about social evaluation [74,75], it suggests that this may be a sensitive period for social norm learning. In collectivist cultures like China, which prioritize group harmony, this sensitivity could underpin the internalization of norms that favor ingroup members. Recognizing this sensitivity underscores the importance of tailored educational strategies. Previous research has demonstrated the effectiveness of bystander interventions in reducing bias-based bullying during adolescence [43,76]. However, adolescents often exhibit restrained prosocial bystander behaviors, and their willingness to intervene against ingroup members' unfair behavior diminishes with age [77–79]. Thus, fostering a social environment that promotes altruistic punishment is crucial [80,81]. Schools can leverage social incentives to cultivate such environments.

A salient feature of our study centers on how individuals' pre-existing social experiences shape reinforcement learning within specific social contexts. In contrast to prior reinforcement learning studies employing probabilistic tasks, where actions arbitrarily corresponded with outcomes (e.g., selecting an image might lead to a reward) [82–84], our paradigm incorporates a socially contextualized learning process where actions bear moral and social significance. This methodology has greater ecological validity, providing insights into the intricate social-cultural learning dynamics. Moreover, it holds potential in shedding light on dysfunctional reward processing, which is central to the pathophysiology of mental disorders like depression [84] and anxiety [85]. Such mental disorders frequently result in socially maladaptive behaviors, including impairments in understanding, following, and enforcing social norms [86]. A probable contributor to these issues is a deficiency in social reward learning. Despite this, the bulk of existing research on reward learning in mental disorders has been confined to cognitive dimensions employing monetary incentives [87]. To achieve a comprehensive understanding of the factors influencing adaptive social decision-making, it is critical to focus on the role of reward learning during social interactions.

The limitations of this study and its findings suggest avenues for further research. First, our study highlights the significant differences in learning altruistic punishment across cultures and ages, even in arbitrarily assigned groups. Research indicates that the ingroup bias observed in minimal groups can predict biases across diverse real-world social groups, including families, friendships, and ethnic communities [57,58]. Nevertheless, the significant variability in social decision-making among these groups underscores the need for further investigation into how social feedback learning influences decision-making within natural groups [57,58]. Second, following common practice in cross-cultural research, our study employs country as a proxy for culture, which may not fully account for confounding variables between the Western and Chinese samples [71]. Despite demographic similarities in our samples, unaddressed socio-ecological factors (e.g., community size, pathogen prevalence) may still influence ingroup bias and norm learning [2,15]. Future research would benefit from accounting for these elements to better disentangle the cultural effects. Third, our findings suggest the

influence of cultural and developmental factors on altruistic behavior. However, the causality between these elements remains unclear. It is recommended that future research should manipulate cultural norms and employ longitudinal methods in order to establish causality and trace the development of these phenomena. Fourth, our study did not include fair trials in order to streamline the learning process. However, future studies should investigate how participants switch their learning of punishment between mixed contexts (e.g., fair vs. unfair conditions) through more complex learning strategies, such as model-based learning [88,89].

In sum, our unique approach allowed us to illuminate the potential causal effect of social feedback on group-biased altruistic punishment, and elucidate the distinct computational learning mechanisms across cultures and age groups. This approach facilitated an exploration of the interplay between experimentally manipulated norms and culturally ingrained norms, which are typically persistent and have deep roots. Our study particularly reveals distinct reinforcement learning processes across different cultures: Chinese adults were overall more responsive to social feedback than their American counterparts, and adjusted their decisions more slowly when punishing ingroup members than outgroups. Notably, this pattern was absent among Americans, indicating a pronounced ingroup bias in collectivist cultures. Furthermore, this ingroup bias is primarily driven by an increasing ingroup favoritism from adolescence to adulthood, rather than outgroup harshness, contributing to the debate over the formation of group bias. Our findings highlight the theoretical significance of how socio-cultural norms impact the underlying learning mechanisms during the social decision-making processes, and provide valuable insights for future studies to assess practical implications and effectiveness in fostering altruistic norm enforcement behaviors in a culturally sensitive manner.

Materials and methods

Ethics statement

Experiments in this study were approved by the Research Ethics Committee, Department of Psychology, Tsinghua University, China (ref: 2021–26). For adult participants, as all tasks were completed online, participants provided their consent by ticking checkboxes after reading the information sheet and prior to the administration of tasks on computers. For child participants, formal written consent was obtained from both the children and their guardians.

Participants

For the power analysis, we utilized the G*Power software package (version 3.1.9.7). In Study 1, consistent with prior research [7,17], we assumed a small-to-medium effect size (Cohen's $f = .14$, $\alpha = .05$, power $[1 - \beta] = 95\%$) to detect the interaction effect of the group membership of dividers and culture, which resulted in a total sample size of 405 participants. Therefore, we initially recruited 440 participants; however, we excluded 51 participants who did not complete the main altruistic punishment task, resulting in a final sample of 389 participants (see Table 1 for the demographic information). Specifically, Chinese participants ($n = 217$, $M_{age} = 28.01$, $SD_{age} = 7.46$; 69 males) were recruited online through Credamo [7], while Western participants ($n = 172$, $M_{age} = 35.01$, $SD_{age} = 9.90$; 89 males) were recruited online through Prolific [90]. Informed consent was obtained from all participants before the commencement of the study. In study 2, to examine how ingroup preferences among Chinese participants develop with age, we recruited 213 adolescent participants ($M_{age} = 15.01$, $SD_{age} = 1.82$, range from 12 to 18; 108 males) and combined their data with the data from 217 adult Chinese participants in Study 1, resulting in a total of 430 participants. The study method was the same as in Study 1.

Table 1. Demographic information of Chinese and American adults in Study 1 and Chinese adolescents Study 2.

	Chinese adults	American adults	Chinese adolescents
Age in years, <i>M</i> (<i>SD</i>)	28.01 (7.46)	35.01 (9.90)	15.98 (1.82)
Gender, <i>n</i> (% Male)	69 (32.8%)	89 (51.7%)	108 (50.7%)
Subjective socioeconomic status, <i>M</i> (<i>SD</i>)	5.18 (2.43)	5.76 (1.93)	5.73 (1.72)
Education level, <i>M</i> (<i>SD</i>)	1.98 (0.90)	2.54 (0.97)	1(0)

<https://doi.org/10.1371/journal.pcbi.1012274.t001>

Altruistic punishment reinforcement learning task

The present study aimed to investigate how individuals dynamically adjust their behavior in response to external feedback, specifically in the context of altruistic punishment. To this end, we employed a commonly used altruistic punishment task and used computational modeling of reinforcement learning to capture participants' trial-by-trial behavior changes. The task consisted of two stages: (1) Punishment Decision-Making and (2) Feedback.

Punishment Decision-Making. In this stage, the game involved three players. Players 1 and 2 each contributed equally and earned a reward of 3 RMB by correctly solving three calculation problems. Player 1 was then given the reward and had the option to allocate half of it (1.5 RMB) to Player 2 or keep it all. Player 3, endowed with 1 RMB, observed Player 1's allocation and could choose to accept or reject it. A rejection by Player 3 would result in both Player 1 and Player 3 receiving 0 RMB, categorizing this action as costly altruistic punishment. If Player 3 accepted the allocation, both Player 1 and Player 3 retained their respective endowments. Additionally, Player 3's choices had no impact on Player 2's endowments; that is, even if Player 3 chose to punish Player 1's unfair allocation, Player 2 did not receive any compensation and only received 0 RMB. This approach aimed to directly assess the effect of punishment, without the influence of compensation motives.

Feedback. Immediate social feedback was provided to Player 3 regarding their decision to punish or accept the allocation. Player 3 was informed that their decisions would be observed by other players in the waiting room, who would provide feedback during each round. Positive social feedback (a thumbs-up) was given when more than 60% of the other players approved of Player 3's choice, while negative feedback (a thumbs-up with a red cross) was given when less than 60% of the other players approved. Participants were randomly assigned to one of the two social norm conditions: the reward-to-punish or reward-to-accept condition. In the reward-to-punish condition, participants were incentivized to punish: punishment resulted in positive feedback with an 80% probability and negative feedback with a 20% probability. In the reward-to-accept condition, participants were incentivized to accept unfair allocations: acceptance resulted in positive feedback with an 80% probability and negative feedback with a 20% probability.

Manipulation of group membership

To examine group bias in third-party punishment (TPP), we manipulated the divider's group membership (ingroup vs. outgroup) using the minimal group paradigm [50,57]. Participants chose either a blue, yellow, or red team based on color preference. This strategy is widely utilized in social psychology research and aids in minimizing the influence of participants' varying previous experiences with actual social groups [50,57]. Furthermore, aligning with prior research [7,17], we ensured Player 2 belonged to a different team from Player 1 and Player 3, making the receiver an outgroup member from the viewpoint of participants.

Procedure

First, participants chose to join a team. Then they were informed that they would be randomly assigned to be Player 1, 2, or 3 and would interact with other players anonymously, and their decisions in the game would affect their payoffs. Indeed, all participants were assigned to be Player 3, and other players, including Players 1 and 2, and observers in the waiting room, were all simulated by programming. Player 1 could be either an ingroup or outgroup member of Player 3. Player 2 was always from a different team than Player 1 and Player 3.

Next, the experiment started and consisted of three phases: pre-test, reinforcement learning, and post-test. In the pre-test phase, participants made decisions without receiving feedback, and were told that players in the waiting rooms did not know their choices. There were 10 trials in the pre-test phase, all of which were unfair allocations, with five trials for each ingroup and outgroup condition. These decisions served as a baseline for altruistic punishment.

Next, participants completed the reinforcement learning phase, in which they received probabilistic feedback after each decision. There were two blocks for each ingroup and outgroup condition, with 15 trials per block, resulting in a total of 60 trials. The order of ingroup and outgroup conditions was randomized across participants.

Finally, in the post-test phase, participants completed ten trials without feedback, including five trials for each ingroup and outgroup condition, to determine whether the effect of the feedback lasted. Besides, we exclusively used unfair allocation scenarios across all trials, for the following two reasons. First, since fair allocations are generally accepted with little variability in responses, they offer limited scopes for learning through social feedback. By focusing on unfair distributions, we aimed to enhance the scope for learning. Second, including fair allocations could require participants to adapt to varying conditions, thereby complicating the learning process with more advanced strategies like model-based reinforcement learning. Therefore, by concentrating solely on unfair scenarios, our experiment avoided these complexities, facilitating a clearer analysis of the responses to unfairness. After completing the game, to provide an index of group identity, participants rated the extent to which they identified to their group on a 7-point scale (“*How much do you feel that you belong to this group?*”, 1 = very low, 7 = very high). Besides, to assess the experiment’s authenticity, participants evaluated their belief in the real-time interaction with other players during the game (“*How much do you believe that you interacted with other players online in real time?*”, 1 = very low, 7 = very high). Then participants reported their demographic information including gender, age, education level (1 = not a high school graduate, 5 = graduate school/postgraduate training) and subjective socioeconomic status (SES). The SES was measured using the social ladder task from 1 (bottom) to 10 (top) [71]. A higher score on the scale indicates that participants perceive their economic status to be more favorable relative to others in society.

Computational modeling

Computational modeling was conducted using MATLAB. As in previous studies, we modeled the data using an R-W reinforcement learning model.

$$V_{t+1}(a) = V_t(a) + \alpha_{i(j)} \cdot PE_t \quad (1)$$

The model estimated how each participant updated the expected values (V) of each choice (punish or accept) based on the trial-by-trial feedback they had received (see Eqs 1 and 2). The V values were initialized at each participant’s probability of choosing each option in the pre-test phase, ranging from 0 to 1. The subscript i in the equation represented the divider group (ingroup vs. outgroup), while subscript j represented the order of blocks (1st vs. 2nd). The learning rates were scaling parameters that adjusted the amplitude of value changes from one

trial to the next.

$$PE_t = R_t - V_t(a) \quad (2)$$

After each trial, the value of the chosen option was updated based on the prediction errors (PE, see Eq 2) of the chosen option. The PEs were computed as the difference between the current reward received ($R = 1$ if positive feedback, $R = 0$ if negative feedback) and the expected value of the chosen option, thus PEs represent the extent to which environmental rewards diverge from expected outcomes.

The softmax link function was utilized to model the relationship between the expected value of a particular action $V_{t(a)}$, and the probability of choosing that action on trial t :

$$P_t(a|V_t(a)) = \frac{e^{(V_t(a)/\beta)}}{\sum_{a'} e^{(V_t(a')/\beta)}} \quad (3)$$

Crucially, the decision calculus is refined by the ingroup bias term $Bias_{(i,a)}$ as introduced into computations through the softmax function which determines the selection probability for each action.

$$P_{i,t}(a|V_t(a)) = \frac{1}{1 + \sum_{a' \neq a} e^{-(V(a) - V_t(a') + Bias_{(i,a)})/\beta}} \quad (4)$$

The bias term, $Bias_{(i,a)}$, is dynamically applied and is contingent on the divider group subscript i and the action subscript a (see Eq 5).

$$Bias_{(i,a)} = \begin{cases} \text{bias} & \text{if } i = \text{ingroup and } a = \text{accept} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Specifically, for the ingroup condition, the bias term is added to the value deviations between the choice of acceptance and punishment, thereby increasing the likelihood of opting for acceptance while simultaneously decreasing the propensity to select punishment. This means that for ingroup members, irrespective of whether the norm is to punish or to accept, the bias specifically applies to increase the tendency to accept ingroup members' behaviors, indicating a potential inclination to follow the ingrained cultural preference. For the outgroup condition, the bias term does not influence the deviation between values for various choices, reflecting an absence of particular cultural bias toward outgroup members.

β represents the temperature. A high β value indicates that choices appear random, with equal likelihood of choosing each option regardless of expected value, while a low β value results in consistently selecting the option with the highest expected value across all trials.

Model fitting and comparison

Models were fitted using a hierarchical approach and compared using Integrated Bayesian Information Criterion (Intergated BIC). We tested multiple models that varied with respect to whether learning could be explained by shared or separate free parameters across group members (ingroup and outgroup dividers) and block numbers (first and second block). We examined whether shared or separate learning rates and temperature, in particular, resulted in a better model fit. Specifically, we compared six candidate models:

- i. $1\alpha 1\beta$: one α and one β for all group dividers and blocks;
- ii. $2\alpha 1\beta$: one α for ingroup dividers and one α for outgroup dividers; one β for all group dividers and blocks;

- iii. $2\alpha 2\beta$: separate α and β for both ingroup and outgroup dividers;
- iv. $4\alpha 1\beta$: separate α for both ingroup and outgroup dividers in the first and second blocks; one β for both ingroup and outgroup dividers;
- v. $4\alpha 2\beta$: separate α for both ingroup and outgroup dividers in the first and second blocks; one β for ingroup dividers and one β for outgroup dividers;
- vi. $4\alpha 2\beta + \text{bias}$: separate α and β for both ingroup and outgroup dividers in the first and second blocks, with a bias term

To examine the association between parameters (learning rate and temperature) and punishment rates for our task, the present study utilized a simulation approach to obtain data from a large sample of 10,000 participants. Specifically, we drew the learning rates (α) and temperature (β) from beta and gamma distributions, respectively. Simulated participants had a distinct learning rate for each divider group (ingroup vs. outgroup) and block (first vs. second), ranging from 0 to 1 to produce 40,000 learning rate values. Moreover, consistent with previous research [83], we evaluated our winning model’s reliability via parameter recovery on simulated data. Simulating choices 15,625 times using our experimental schedule and employing an iterative maximum a posteriori (MAP) approach fitting revealed strong correlations (see Fig 3A) between simulated and fitted parameter values, indicating effective model parameter estimation in our experiment.

The MAP approach offers improvements over traditional maximum likelihood estimation (MLE) through a hierarchical analysis comprising individual subject assessments and collective sample evaluations. Initially, we set uninformative priors with means $\mu = 0.1$ (plus noise) and a variance $\sigma^2 = 100$ the group-level Gaussian distributions. First, during the expectation step, we calculated the log-likelihood of the anticipation’s series of choice given a model M and its parameter vector $\eta_i(\alpha \text{ and } \beta)$ for each participant i ($i \in [1, N]$). We summed the conditional probability of each trial’s choice given the model’s parameter η_i and initial expectation of values for both punishment and non-punishment $V'_i = [V_a; V_b]$ over all trials. The prior probability of each participant’s η_i was calculated given the group-level Gaussian distributions over the parameters (with mean of μ and variance of σ^2). Given the probability of the already observed choice series is constant, we computed the posterior probability ($P_{\text{posterior},i}$) estimate as follows:

$$P_{\text{posterior},i} = \max_{\eta_i} \left[\sum_t (\log(p_t(\text{choice}_t | \eta_i, V'_i)) + \log(\text{normpdf}(\eta_i | \mu, \sigma^2))) \right]$$

$$\eta_i = \text{argmax}_{\eta_i} \left[\sum_t (\log(p_t(\text{choice}_t | \eta_i, V'_i)) + \log(\text{normpdf}(\eta_i | \mu, \sigma^2))) \right]$$

Second, during the maximization step, we recomputed μ and σ^2 based on the estimated set of η and their Hessian matrix H (as calculated with Matlab’s `fminunc`) over all N participants.

$$\mu = \frac{1}{N} \sum_i \eta_i$$

$$\sigma^2 = \frac{1}{N} \sum_i [\eta_i^2 + \text{diag}(\text{pinv}(H_i))] - \mu^2$$

Where the diagonal terms of the inverted Hessian matrix (computed in Matlab with `diag(pinv(Hi))`) give the second moment around η_i , approximating the variance, and thus the inverse of

the uncertainty with which the parameter can be estimated. We repeated expectation and maximization step until convergence—defined as a posterior likelihood change under 0.001 between iterations—or a maximum of 800 iterations were completed. Free parameters that were bounded underwent careful transformation using link functions, such as sigmoid transformations for learning rate parameters, to maintain estimate precision.

Statistical analysis

We used R Version 4.1.0 (80) to conduct all statistical analyses on behavioral data and computational modeling parameters. Models were all run using the package ‘lme4’. We employed generalized linear mixed models (GLMM) to analyze the decisions in altruistic punishment game due to binary responses (punish = 1; accept = 0), and we used linear mixed-effects models (LMM) to analyze the continuous learning rates (α). In Study 1, fixed terms included cultural background, dividers’ membership and punishment norms. Age, gender, education level, and subjective socioeconomic status were included as covariates to control potential confounding demographic differences between American and Chinese samples. In Study 2, fixed terms included age, dividers’ membership and punishment norm. Gender, education level, and subjective socioeconomic status were included as covariates. The model selection process began with the full model, which incorporated random intercepts for each participant and random slopes for all fixed within-group factors [91]. We then fitted a series of reduced models, each excluding one of the random slopes, and selected the model with the lowest Akaike Information Criterion (AIC). This model was subsequently compared to the more complex original using a likelihood-ratio test. We opted for the more complex model if the p-value from the χ^2 -statistic was less than 0.2 [92]. If not, we continued to simplify the model until an optimal one was chosen or all random slopes were removed. Models that failed to converge were excluded from this analysis. For two-level factors, we applied sum contrasts—for example, for group membership (ingroups = -0.5, outgroups = 0.5), cultural background (Chinese = -0.5, American = 0.5), norms regarding punishment (punishment-encouragement = -0.5, acceptance-encouragement = 0.5), order of blocks (first block = -0.5, second block = 0.5), and gender (male = -0.5, female = 0.5). Additionally, we centered all continuous predictors, such as age, education level, and subjective socioeconomic status, around their mean values. Data and analysis scripts can be obtained online (<https://doi.org/10.5281/zenodo.11239516>).

Supporting information

S1 Table. Model comparison and the model selection process for punishment behaviors in pre-test stage in Study 1.

(DOC)

S2 Table. Model results for punishment behaviors in pre-test stage in Study 1.

(DOC)

S3 Table. Model comparison and the model selection process for punishment behaviors in learning stage in Study 1.

(DOC)

S4 Table. Model results for punishment behaviors in learning stage in Study 1.

(DOC)

S5 Table. Study 1 Model Comparison Results.

(DOC)

S6 Table. Parameter Recovery Results.

(DOC)

S7 Table. Model comparison and the model selection process for learning rates in Study 1.

(DOC)

S8 Table. Model results for learning rates in Study 1.

(DOC)

S9 Table. Model comparison and the model selection process for temperature in Study 1.

(DOC)

S10 Table. Model results for temperature in Study 1.

(DOC)

S11 Table. Model comparison and the model selection process for punishment behaviors in pre-test stage in Study 2.

(DOC)

S12 Table. Model results for punishment behaviors in pre-test stage in Study 2.

(DOC)

S13 Table. Model comparison and the model selection process for punishment behaviors in learning stage in Study 2.

(DOC)

S14 Table. Model results for punishment behaviors in learning stage in Study 2.

(DOC)

S15 Table. Study 2 Model Comparison Results.

(DOC)

S16 Table. Model comparison and the model selection process for learning rates in Study 2.

(DOC)

S17 Table. Model results for learning rates in Study 2.

(DOC)

S18 Table. Model comparison and the model selection process for adolescents' learning rates in Study 2.

(DOC)

S19 Table. Model results for adolescents' learning rates in Study 2.

(DOC)

S20 Table. Model comparison and the model selection process for temperature in Study 2.

(DOC)

S21 Table. Model results for temperature in Study 2.

(DOC)

Author Contributions**Conceptualization:** Ziyang Guo, Zhen Wu.**Data curation:** Ziyang Guo, Jialu Yu.

Formal analysis: Ziyao Guo.

Funding acquisition: Zhen Wu.

Investigation: Ziyao Guo, Wenxin Wang.

Methodology: Ziyao Guo, Jialu Yu, Patricia Lockwood, Zhen Wu.

Project administration: Wenxin Wang.

Resources: Zhen Wu.

Software: Ziyao Guo, Jialu Yu.

Supervision: Zhen Wu.

Validation: Ziyao Guo.

Visualization: Ziyao Guo.

Writing – original draft: Ziyao Guo, Zhen Wu.

Writing – review & editing: Ziyao Guo, Patricia Lockwood, Zhen Wu.

References

1. 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](https://doi.org/10.1016/S1090-5138(04)00005-4)
2. Henrich J, Ensminger J, McElreath R, Barr A, Barrett C, Bolyanatz A, et al. Markets, religion, community size, and the evolution of fairness and punishment. *Science.* 2010 Mar 19; 327 (5972):1480–4. <https://doi.org/10.1126/science.1182238> PMID: 20299588
3. Henrich J, McElreath R, Barr A, Ensminger J, Barrett C, Bolyanatz A, et al. Costly punishment across human societies. *Science.* 2006 Jun 23; 312(5781):1767–70. <https://doi.org/10.1126/science.1127333> PMID: 16794075
4. Jordan JJ, Hoffman M, Bloom P, Rand DG. Third-party punishment as a costly signal of trustworthiness. *Nature.* 2016 Feb 25; 530(7591):473–6. <https://doi.org/10.1038/nature16981> PMID: 26911783
5. Bernhard H, Fischbacher U, Fehr E. Parochial altruism in humans. *Nature.* 2006 Aug 24; 442 (7105):912–5. <https://doi.org/10.1038/nature04981> PMID: 16929297
6. Delton AW, Krasnow MM. The psychology of deterrence explains why group membership matters for third-party punishment. *Evol Hum Behav.* 2017; 38:734–743. <https://doi.org/10.1016/j.evolhumbehav.2017.07.003>
7. Guo Z, Guo R, Xu C, Wu Z. 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>
8. McAuliffe K, Dunham Y. Group bias in cooperative norm enforcement. *Philos Trans R Soc Lond B Biol Sci.* 2016 Jan 19; 371(1686):20150073. <https://doi.org/10.1098/rstb.2015.0073> PMID: 26644592
9. Marshall J, McAuliffe K. Children as assessors and agents of third-party punishment. *Nat Rev Psychol.* 2022; 1:334–344. <https://doi.org/10.1038/S44159-022-00046-Y>
10. Amir D, McAuliffe K. Cross-cultural, developmental psychology: integrating approaches and key insights. *Evol Hum Behav.* 2020; 41:430–444. <https://doi.org/10.1016/j.evolhumbehav.2020.06.006>
11. Chudek M, Henrich J. Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends Cogn Sci.* 2011 May; 15(5):218–26. <https://doi.org/10.1016/j.tics.2011.03.003> PMID: 21482176
12. Bicchieri C. *The grammar of society: The nature and dynamics of social norms.* Cambridge University Press; 2006.
13. Cialdini RB, Wosinska W, Barrett DW, Butner J, Gornik-Durose M. Compliance with a request in two cultures: the differential influence of social proof and commitment/consistency on collectivists and individualists. *Pers Soc Psychol Bull.* 1999; 25:1242–1253. <https://doi.org/10.1177/0146167299258006>
14. Fehr E, Schurtenberger I. Normative foundations of human cooperation. *Nat Hum Behav.* 2018 Jul; 2 (7):458–468. <https://doi.org/10.1038/s41562-018-0385-5> PMID: 31097815
15. Hruschka DJ, Henrich J. Economic and evolutionary hypotheses for cross-population variation in parochialism. *Front Hum Neurosci.* 2013 Sep 11; 7:559. <https://doi.org/10.3389/fnhum.2013.00559> PMID: 24062662

16. Yang F, Yang X, Dunham Y. Beyond our tribe: Developing a normative sense of group-transcendent fairness. *Dev Psychol*. 2023 Jul; 59(7):1203–1217. <https://doi.org/10.1037/dev0001545> PMID: 37166870
17. Yudkin DA, Rothmund T, Twardawski M, Thalla N, Van Bavel JJ. Reflexive intergroup bias in third-party punishment. *J Exp Psychol Gen*. 2016 Nov; 145(11):1448–1459. <https://doi.org/10.1037/xge0000190> PMID: 27632379
18. Graham J, Meindl P, Beall E, Johnson KM, Zhang L. Cultural differences in moral judgment and behavior, across and within societies. *Curr Opin Psychol*. 2016 Apr; 8:125–130. <https://doi.org/10.1016/j.copsyc.2015.09.007> PMID: 29506787
19. Markus HR, Kitayama S. Culture and the self: implications for cognition, emotion, and motivation. *Psychol Rev*. 1991; 98(2):224–253. <https://doi.org/10.1037//0033-295x.98.2.224>
20. Talhelm T, Zhang X, Oishi S, Shimin C, Duan D, Lan X, et al. Large-scale psychological differences within china explained by rice versus wheat agriculture. *Science*. 2014 May 9; 344(6184):603–8. <https://doi.org/10.1126/science.1246850> PMID: 24812395
21. Triandis HC. Differing Cultural Contexts. *The culture and psychology reader*. 1995; 98: 326.
22. House BR, Kanngiesser P, Barrett HC, Yilmaz S, Smith AM, Sebastián-Enesco C, et al. Social norms and cultural diversity in the development of third-party punishment. *Proc Biol Sci*. 2020 Apr 29; 287(1925):20192794. <https://doi.org/10.1098/rspb.2019.2794> PMID: 32315587
23. Morris MW, Savani K, Fincher K. Metacognition fosters cultural learning: Evidence from individual differences and situational prompts. *J Pers Soc Psychol*. 2019 Jan; 116(1):46–68. <https://doi.org/10.1037/pspi0000149> PMID: 30596446
24. Savani K, Morris MW, Fincher K, Lu JG, Kaufman SB. Experiential learning of cultural norms: The role of implicit and explicit aptitudes. *J Pers Soc Psychol*. 2022 Aug; 123(2):272–291. <https://doi.org/10.1037/pspa0000290> PMID: 35099201
25. House BR, Kanngiesser P, Barrett HC, Broesch T, Crittenden AN, Erut A, et al. Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nat Hum Behav*. 2020 Jan; 4(1):36–44. <https://doi.org/10.1038/s41562-019-0734-z> PMID: 31548679
26. Balliet D, Wu J, De Dreu CK. Ingroup favoritism in cooperation: a meta-analysis. *Psychol Bull*. 2014 Nov; 140(6):1556–1581. <https://doi.org/10.1037/a0037737> PMID: 25222635
27. Iacoviello V, Spears R. “I know you expect me to favor my ingroup”: reviving Tajfel’s original hypothesis on the generic norm explanation of ingroup favoritism. *J Exp Soc Psychol*. 2018; 76:88–99. <https://doi.org/10.1016/j.jesp.2018.01.002>
28. Triandis HC, Bontempo R, Villareal MJ, Asai M, Lucca N. Individualism and collectivism: cross-cultural perspectives on self-ingroup relationships. *J Pers Soc Psychol*. 1988; 54:323–338. <https://doi.org/10.1037/0022-3514.54.2.323>
29. Miller JG, Das R, Chakravarthy S. Culture and the role of choice in agency. *J Pers Soc Psychol*. 2011 Jul; 101(1):46–61. <https://doi.org/10.1037/a0023330> PMID: 21480735
30. Trommsdorff G, Chen X. Values, religion, and culture in adolescent development. Cambridge: Cambridge University Press; 2012.1–448p.
31. Gelfand MJ, Gavrilets S, Nunn N. Norm Dynamics: Interdisciplinary perspectives on social norm emergence, persistence, and change. *Annu Rev Psychol*. 2024 Jan 18; 75:341–378. <https://doi.org/10.1146/annurev-psych-033020-013319> PMID: 37906949
32. Mu Y, Kitayama S, Han S, Gelfand MJ. How culture gets embrained: cultural differences in event-related potentials of social norm violations. *Proc Natl Acad Sci U S A*. 2015 Dec 15; 112(50):15348–15353. <https://doi.org/10.1073/PNAS.1509839112> PMID: 26621713
33. Stamkou E, van Kleef GA, Homan AC, Gelfand MJ, van de Vijver FJR, van Egmond MC, et al. Cultural collectivism and tightness moderate responses to norm violators: effects on power perception, moral emotions, and leader support. *Pers Soc Psychol Bull*. 2019 Jun; 45(6):947–964. <https://doi.org/10.1177/0146167218802832> PMID: 30394858
34. Kish Bar-On K, Lamm E. The interplay of social identity and norm psychology in the evolution of human groups. *Philos Trans R Soc Lond B Biol Sci*. 2023 Mar 13; 378(1872):20210412. <https://doi.org/10.1098/rstb.2021.0412> PMID: 36688389
35. Shalvi S, Handgraaf MJJ, De Dreu CKW. People avoid situations that enable them to deceive others. *J Exp Soc Psychol*. 2011; 47:1096–1106. <https://doi.org/10.1016/j.jesp.2011.04.015>
36. Will GJ, Rutledge RB, Moutoussis M, Dolan RJ. Neural and computational processes underlying dynamic changes in self-esteem. *ELife*. 2017 Oct 24; 6:e28098. <https://doi.org/10.7554/eLife.28098> PMID: 29061228

37. van Baar JM, Nassar MR, Deng W, FeldmanHall O. Latent motives guide structure learning during adaptive social choice. *Nat Hum Behav.* 2022 Mar; 6(3):404–414. <https://doi.org/10.1038/s41562-021-01207-4> PMID: 34750584
38. Ciranka S, Linde-Domingo J, Padezhki I, Wicharz C, Wu CM, Spitzer B. Asymmetric reinforcement learning facilitates human inference of transitive relations. *Nat Hum Behav.* 2022 Apr; 6(4):555–564. <https://doi.org/10.1038/s41562-021-01263-w> PMID: 35102348
39. Heffner J, Son JY, FeldmanHall O. Emotion prediction errors guide socially adaptive behaviour. *Nat Hum Behav.* 2021 Oct; 5(10):1391–1401. <https://doi.org/10.1038/s41562-021-01213-6> PMID: 34667302
40. Zaki J, Kallman S, Wimmer GE, Ochsner K, Shohamy D. Social cognition as reinforcement learning: feedback modulates emotion inference. *J Cogn Neurosci.* 2016 Sep; 28(9):1270–1282. https://doi.org/10.1162/jocn_a_00978 PMID: 27167401
41. Zhang L, Lengersdorff L, Mikus N, Gläscher J, Lamm C. Using reinforcement learning models in social neuroscience: frameworks, pitfalls and suggestions of best practices. *Soc Cogn Affect Neur.* 2020 Jul 30; 15(6): 695–707. <https://doi.org/10.1093/scan/nsaa089> PMID: 32608484
42. Pauli R, Brazil IA, Kohls G, Klein-Flügge MC, Rogers JC, Dikeos D, et al. Action initiation and punishment learning differ from childhood to adolescence while reward learning remains stable. *Nat Commun.* 2023 Sep 14; 14(1):5689. <https://doi.org/10.1038/s41467-023-41124-w> PMID: 37709750
43. Elenbaas L, Rizzo MT, Killen M. A developmental-science perspective on social inequality. *Curr Dir Psychol Sci.* 2020 Dec 1; 29(6):610–616. <https://doi.org/10.1177/0963721420964147> PMID: 33758480
44. Jordan JJ, McAuliffe K, Warneken F. Development of in-group favoritism in children's third-party punishment of selfishness. *Proc Natl Acad Sci U S A.* 2014 Sep 2; 111(35):12710–5. <https://doi.org/10.1073/pnas.1402280111> PMID: 25136086
45. Engelmann JM, Herrmann E, Rapp DJ, Tomasello M. Young children (sometimes) do the right thing even when their peers do not. *Cogn Dev.* 2016; 39:86–92. <https://doi.org/10.1016/j.cogdev.2016.04.004>
46. Helwig CC, Jasiobedzka U. The relation between law and morality: children's reasoning about socially beneficial and unjust laws. *Child Dev.* 2001 Sep-Oct; 72(5): 1382–93. <https://doi.org/10.1111/1467-8624.00354> PMID: 11699676
47. Elenbaas L, Killen M. How do young children expect others to address resource inequalities between groups? *J Exp Child Psychol.* 2016 Oct; 150:72–86. <https://doi.org/10.1016/j.jecp.2016.05.002> PMID: 27262524
48. Tomasello M. *Becoming human: a theory of ontogeny.* Cambridge: Harvard University Press; 2019.
49. Hogg MA, Sherman DK, Dierselhuis J, Maitner AT, Moffitt G. Uncertainty, entitativity, and group identification. *J Exp Soc Psychol.* 2007; 43:135–142. <https://doi.org/10.1016/j.jesp.2005.12.008>
50. Tajfel H, Billig MG, Bundy RP, Flament C. Social categorization and intergroup behaviour. *Eur J Soc Psychol.* 1971; 1:149–178. <https://doi.org/10.1002/ejsp.2420010202>
51. Romano A, Sutter M, Liu JH, Yamagishi T, Balliet D. National parochialism is ubiquitous across 42 nations around the world. *Nat Commun.* 2021 Jul 22; 12(1):4456. <https://doi.org/10.1038/s41467-021-24787-1> PMID: 34294708
52. Pickup M, Kimbrough EO, De Rooij EA. Expressive politics as (costly) norm following. *Polit Behav.* 2022; 44:1611–1631. <https://doi.org/10.1007/s11109-020-09667-6>
53. Packer DJ, Chasteen AL. Loyal deviance: testing the normative conflict model of dissent in social groups. *Pers Soc Psychol Bull.* 2010 Jan; 36(1):5–18. <https://doi.org/10.1177/0146167209350628> PMID: 19907038
54. Ellemers N, Jetten J. The many ways to be marginal in a group. *Pers Soc Psychol Rev.* 2013 Feb; 17(1):3–21. <https://doi.org/10.1177/1088868312453086> PMID: 22854860
55. Masson T, Fritsche I. Loyal peripherals? The interactive effects of identification and peripheral group membership on deviance from non-beneficial ingroup norms. *Eur J Soc Psychol.* 2019; 49:76–92. <https://doi.org/10.1002/ejsp.2501>
56. Gomila R, Paluck EL. The social and psychological characteristics of norm deviants: a field study in a small cohesive university campus. *J Soc Polit Psychol.* 2020 Feb 28; 8(1):220–45. <https://doi.org/10.5964/jsp.v8i1.1134>
57. Dunham Y. Mere Membership. *Trends Cogn Sci.* 2018 Sep; 22(9):780–793. <https://doi.org/10.1016/j.tics.2018.06.004> PMID: 30119749
58. Akrami N, Ekehammar B, Bergh R. Generalized prejudice: common and specific components. *Psychol Sci.* 2011 Jan; 22(1):57–9. <https://doi.org/10.1177/0956797610390384> PMID: 21106890

59. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci*. 2007 Sep; 10(9):1214–21. <https://doi.org/10.1038/nn1954> PMID: 17676057
60. Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci U S A*. 2007 Oct 9; 104(41):16311–6. <https://doi.org/10.1073/pnas.0706111104> PMID: 17913879
61. Soltani A, Izquierdo A. Adaptive learning under expected and unexpected uncertainty. *Nat Rev Neurosci*. 2019 Oct; 20(10):635–644. <https://doi.org/10.1038/s41583-019-0180-y> PMID: 31147631
62. Schiller B, Baumgartner T, Knoch D. Intergroup bias in third-party punishment stems from both ingroup favoritism and outgroup discrimination. *Evolution and Human Behavior*. 2014; 35:169–175. <https://doi.org/10.1016/j.evolhumbehav.2013.12.006>
63. Li X, Molleman L, van Dolder D. Do descriptive social norms drive peer punishment? Conditional punishment strategies and their impact on cooperation. *Evol Hum Behav*. 2021; 42: 469–479. <https://doi.org/10.1016/j.evolhumbehav.2021.04.002>
64. House BR. How do social norms influence prosocial development? *Curr Opin Psychol*. 2018 Apr; 20:87–91. <https://doi.org/10.1016/j.copsyc.2017.08.011> PMID: 28858771
65. Morris MW, Hong Y, Chiu C, Liu Z. Normology: Integrating insights about social norms to understand cultural dynamics. *Organ Behav Hum Decis Process*. 2015; 129:1–13. <https://doi.org/10.1016/j.obhdp.2015.03.001>
66. Henrich J. *The Secret of Our Success: How culture is driving human evolution, domesticating our Species, and making us smarter*. Princeton: Princeton University Press; 2016. <https://doi.org/10.2307/j.ctvc77f0d>
67. Stubbersfield JM. Content biases in three phases of cultural transmission: a review. *Cult Evol*. 2022; 19:41–60. <https://doi.org/10.1556/2055.2022.00024>
68. Liu SS, Shteynberg G, Morris MW, Yang Q, Galinsky AD. How does collectivism affect social interactions? A test of two competing accounts. *Pers Soc Psychol Bull*. 2021 Mar; 47(3):362–376. <https://doi.org/10.1177/0146167220923230> PMID: 32515282
69. Oyserman D, Coon HM, Kemmelmeier M. Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychol Bull*. 2002 Jan; 128(1):3–72. <https://doi.org/10.1037/0033-2909.128.1.3> PMID: 11843547
70. Yamagishi T, Jin N, Miller AS. In-group bias and culture of collectivism. *Asian J Soc Psychol*. 1998; 1:315–328. <https://doi.org/10.1111/1467-839X.00020>
71. Liu SS, Morris MW, Talhelm T, Yang Q. Ingroup vigilance in collectivistic cultures. *Proc Natl Acad Sci U S A*. 2019 Jul 16; 116(29):14538–14546. <https://doi.org/10.1073/pnas.1817588116> PMID: 31249140
72. Tomasello M. *Why we cooperate*. Cambridge, MA, US: MIT Press; 2009. pp. xviii, 206.
73. O'Madagain C, Tomasello M. Shared intentionality, reason-giving and the evolution of human culture. *Philos Trans R Soc Lond B Biol Sci*. 2022 Jan 31; 377(1843):20200320. <https://doi.org/10.1098/rstb.2020.0320> PMID: 34894741
74. van den Bos E, de Rooij M, Miers AC, Bokhorst CL, Westenberg PM. Adolescents' increasing stress response to social evaluation: pubertal effects on cortisol and alpha-amylase during public speaking. *Child Dev*. 2014 Jan-Feb; 85(1):220–36. <https://doi.org/10.1111/cdev.12118> PMID: 23638912
75. van den Bos E, van Duijvenvoorde AC, Westenberg PM. Effects of adolescent sociocognitive development on the cortisol response to social evaluation. *Dev Psychol*. 2016 Jul; 52(7):1151–1163. <https://doi.org/10.1037/dev0000133> PMID: 27177160
76. Killen M, Dahl A. Moral reasoning enables developmental and societal change. *Perspect Psychol Sci*. 2021 Nov; 16(6):1209–1225. <https://doi.org/10.1177/1745691620964076> PMID: 33621472
77. Mulvey KL, Palmer SB, Abrams D. Race-based humor and peer group dynamics in adolescence: bystander intervention and social exclusion. *Child Dev*. 2016 Sep; 87(5):1379–91. <https://doi.org/10.1111/cdev.12600> PMID: 27684393
78. Palmer SB, Abbott N. Bystander responses to bias-based bullying in schools: a developmental intergroup approach. *Child Dev Perspect*. 2018; 12:39–44. <https://doi.org/10.1111/cdep.12253>
79. Salmivalli C, Voeten M, Poskiparta E. Bystanders matter: associations between reinforcing, defending, and the frequency of bullying behavior in classrooms. *J Clin Child Adolesc Psychol*. 2011; 40(5):668–76. <https://doi.org/10.1080/15374416.2011.597090> PMID: 21916686
80. Nesdale D, Lawson MJ. Social groups and children's intergroup attitudes: can school norms moderate the effects of social group norms? *Child Dev*. 2011 Sep-Oct; 82(5):1594–606. <https://doi.org/10.1111/j.1467-8624.2011.01637.x> PMID: 21883158

81. Yüksel AŞ, Palmer SB, Rutland A. Developmental differences in bystander behavior toward intergroup and intragroup exclusion. *Dev Psychol.* 2021 Aug; 57(8):1342–1349. <https://doi.org/10.1037/dev0001202> PMID: 34591576
82. Cocco VM, Bisagno E, Visintin EP, Cadamuro A, Di Bernardo GA, Trifiletti E, et al. Fighting stigma-based bullying in primary school children: an experimental intervention using vicarious intergroup contact and social norms. *Soc Dev.* 2022; 31:782–796. <https://doi.org/10.1111/sode.12574>
83. Cutler J, Wittmann MK, Abdurahman A, Hargitai LD, Drew D, Husain M, et al. Ageing is associated with disrupted reinforcement learning whilst learning to help others is preserved. *Nat Commun.* 2021 Jul 21; 12(1):4440. <https://doi.org/10.1038/s41467-021-24576-w> PMID: 34290236
84. Hertz U. Learning how to behave: cognitive learning processes account for asymmetries in adaptation to social norms. *Proc Biol Sci.* 2021 Jun 9; 288(1952):20210293. <https://doi.org/10.1098/rspb.2021.0293> PMID: 34074119
85. Frey A-L, Frank MJ, McCabe C. Social reinforcement learning as a predictor of real-life experiences in individuals with high and low depressive symptomatology. *Psychol Med.* 2021 Feb; 51(3):408–415. <https://doi.org/10.1017/S0033291719003222> PMID: 31831095
86. Kupferberg A, Bicks L, Hasler G. Social functioning in major depressive disorder. *Neurosci Biobehav Rev.* 2016 Oct; 69:313–32. <https://doi.org/10.1016/j.neubiorev.2016.07.002> PMID: 27395342
87. Rizvi SJ, Pizzagalli DA, Sproule BA, Kennedy SH. Assessing anhedonia in depression: potentials and pitfalls. *Neurosci Biobehav Rev.* 2016 Jun; 65:21–35. <https://doi.org/10.1016/j.neubiorev.2016.03.004> PMID: 26959336
88. Kool W, Cushman FA, Gershman SJ. When does model-based control pay off? *PLoS Comput Biol.* 2016 Aug 26; 12(8):e1005090. <https://doi.org/10.1371/journal.pcbi.1005090> PMID: 27564094
89. Kurdi B, Gershman SJ, Banaji MR. Model-free and model-based learning processes in the updating of explicit and implicit evaluations. *Proc Natl Acad Sci U S A.* 2019 Mar 26; 116(13):6035–6044. <https://doi.org/10.1073/pnas.1820238116> PMID: 30862738
90. Palan S, Schitter C. Prolific.ac—A subject pool for online experiments. *J Behav Exp Financ.* 2018; 17:22–27. <https://doi.org/10.1016/j.jbef.2017.12.004>
91. Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: keep it maximal. *J Mem Lang.* 2013 Apr; 68(3):10.1016/j.jml.2012.11.001. <https://doi.org/10.1016/j.jml.2012.11.001> PMID: 24403724
92. Matuschek H, Kliegl R, Vasishth S, Baayen H, Bates D. Balancing Type I error and power in linear mixed models. *J Mem Lang.* 2017; 94:305–315. <https://doi.org/10.1016/j.jml.2017.01.001>