Title: The selfish nature in interpersonal exchange among adolescents: imposing the kind and submitting to the exploitative

Abbreviated title: social learning process of adolescents

Author names and affiliations:

Xiaoyan Wu[1,Φ], Chunliang Feng[5,6,Φ], Shen Zhang[1], Zilu Liang[1], Wenshan Dong[5], Shaozheng Qin[1], Yue-Jia Luo[1,2,3,4*] and Chao Liu[1*]

[1] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

[2] Center of Brain Disorder and Cognitive Sciences, College of Psychology and Sociology, Shenzhen University, Shenzhen, China,

[3] Shenzhen Key Laboratory of Affective and Social Cognitive Science, Shenzhen University, Shenzhen, China,

[4] Center for Emotion and Brain, Shenzhen Institute of Neuroscience, Shenzhen, China

[5] Guangdong Provincial Key Laboratory of Mental Health and Cognitive Science, Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, China

[6] Key Laboratory of Brain, Cognition and Education Sciences (South China Normal University), Ministry of Education

***Corresponding authors:**

Yue-Jia Luo, Ph.D.

State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

E-mail: luoyj@szu.edu.cn;

Chao Liu, Ph.D.

State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal

University, Beijing, China

E-mail: liuchao@bnu.edu.cn;

Conflict of Interest: None of the authors have conflicts of interests to declare.

Φ These authors contribute equally to this work.

Abstract

Bullying events during adolescence are common, yet it remains largely unclear which kinds of behavioral patterns are more likely to result in being exploited. Leveraging recent advances in computational modeling and experimental economics, this study examines whether being nice or tough during early social interactions will impact future interactions. Adolescents were randomly assigned to play a cooperation game with two different simulated partners. We found that participants were more likely to cooperate with the initially tough partner than with the initially nice partner. Computational modeling revealed that behavioral changes were driven by changes in perceived social rewards from reciprocity. Perceived social rewards mediated the effects of different partners on the participants' cooperative behaviors. The results indicate that being nice is not a good strategy for building social cooperation, and advance our knowledge of how adolescents form sustained social relationships with peers and may have implications for the education field.

Keywords: social cooperation, adolescents, computational modeling, reinforcement learning, repeated prisoner's dilemma game

Statement of Relevance

Bullying events in high school have been extremely common across cultures. Most studies focus on the modulations of adolescent's behaviors by a specific factor, e.g., education, parenting style or psychological distorter. Our study examines a more general question: whether and how the way of adolescents exhibit themselves would dynamically influence the way that other people treat them. Our findings suggest that being nice initially is not a good strategy for building social cooperation. Being nice early in a relationship will weaken the value of reciprocity perceived by others, which in turn results in being exploited. We regard our work as being of significance for the fields of psychological science, education and human morality. This study advances our knowledge of how adolescents form sustained social relationships with their peers, and it may have implications for the education field.

Introduction

In the development of a social relationship, people calculate the potential risks and benefits from interpersonal interactions [1]. A Chinese proverb says to “try peaceful means before resorting to force” (先禮後兵), which suggests that we should be polite and try to use fair rules in early social relationships. On the other hand, according to another saying, “if you make an ass of yourself, don't complain if people ride you” (人善被人欺), suggesting that we should not be too nice in early interactions; otherwise, we will be exploited. Thus, what is the better strategy for establishing a cooperative relationship with others?

Based on conditional cooperation theory, decisions to engage in cooperation can be strongly affected by the engagement of others [2]. Empirical studies frequently demonstrate that the majority of the population is more likely to contribute more to a team when other players are cooperative [3]. Thus, people being nice will drive cooperation and popularity within social circles, and will be less likely to be exploited. On the other hand, evolutionary theories of social competition propose that people are motivated to increase their personal advantages for fighting over scarce resources due to the survival of the fittest [4]. Indeed, self-comparison with others who are worse off (downward comparison) can elevate people’s positive feelings [5], which can be attributed to the desire to evaluate and enhance self-value by comparing with others [6]. In both humans and primates, decisions that benefit others significantly decrease when the context involves self-interest, suggesting that individuals weight more on their own wellbeing over others [7,8]. In this case, those who are nice are likely to be exploited by the opponent. Therefore, it is unclear whether being nice is a good strategy for establishing mutually cooperative relationships with others without being exploited.

Bullying and victimization are among the most common issues with regard to being exploited in interpersonal interactions during adolescence [9]. Bullying events most frequently occur when children are of high school age [10]. The prevalence of bullying phenomena suggests that adolescents are experiencing setbacks in forming

social cooperation [11]. Social neuroscience proposed that adolescence is a sensitive period for adaptive socioemotional development [12]. In this scenario, what is the best strategy for a teenager to adopt to avoid being exploited in contexts involving cooperation and competition——being nice or being tough? As we know, initial interactions are important for impression formation and further cooperation, and they send a message about one's attitude toward the relationship [13]. Accordingly, information acquired early can be weighted more than that acquired late during impression formation [14]. Here, we examine whether and how being nice or being tough during early social interactions will impact future interactions.

Specifically, 127 adolescents were recruited in the present study to complete a modified version of the prisoner's dilemma game (PDG) with a putative partner (Fig. 1) who was believed to be a real person playing games with the participants in the same lab room. This modified PDG simulated a real social dilemma embedded within cooperation and competition [15]. In the game, the cooperation probabilities of the putative partner were manipulated to simulate two behavioral patterns: being nice (stable cooperation probability) and being tough/erratic (volatile cooperation probability). Moreover, two shift directions were created to simulate different types of partners. One partner behaved nicely at the beginning of the game and then became erratic (stable-volatile); in contrast, regarding the shift direction, the other partner behaved in the opposite manner (volatile-stable). The average cooperation probability remains the same between the two types of partners across the task, and the participants were randomly assigned to one of two groups to play the PDG with one of the simulated partners. This setting allowed us to examine whether and how the early behavioral patterns of an adolescent would impact subsequent social interactions.

Previous PDG-based studies have shown that people significantly decrease their cooperative decisions if they are betrayed by their partner early on compared to during late interactions [13]. These findings are in line with the rule of "trying peaceful means before resorting to force", and they are also consistent with early findings regarding impression formation (for a review, see [16]). In this regard, we expected a

lower cooperation probability in participants who played games with a partner exhibiting a pattern of volatile-stable cooperation. On the other hand, other studies [17] suggest that an early, as opposed to a late, betrayal in an interpersonal relationship leads to low expectations, which in turn make later positive interactions more accepted and cherished. For example, people update their moral impression of a bad person more quickly than they update their moral impression of a good person [18]. These findings imply the opposite hypothesis: participants will cooperate more with a partner exhibiting a pattern of volatile-stable cooperation compared to a partner exhibiting a pattern of stable-volatile cooperation.

Notably, we employed a computational modeling approach to decompose the distinct computational subcomponents underlying the impact of social contexts (i.e., being nice or being tough during early social interactions) on participants' cooperative behaviors. The social value model assumed that participants make decisions by i) updating their expectations about partner cooperativeness under the RL rule as prior work found that adolescents' behavior is captured better by a basic RL algorithm than by other learning modules [19], and ii) receiving social rewards from positive interactions that facilitate future collaborations as many studies have demonstrated the rewarding nature of cooperation itself [20]. The decision theory framework and the algorithm of this model were based on previous work on cooperation [21-23]. Computational modeling allows us to investigate the cognitive subcomponents underlying adolescents' cooperative behaviors.

Results

Being nice during early social interaction results in being exploited by others. Conducting shift direction (between-subjects factor: stable-volatile vs. volatile-stable) × block (within-subjects factor: stable vs. volatile) analyses of variance (ANOVAs) on the cooperation rate with Bonferroni-corrected post hoc t-tests, we found a significant main effect of block (F (1, 125) = 139.128, $p < 0.001$, $\eta^2 = 0.527$): the participants were more likely to cooperate with a partner in the stable block than in the volatile block ($p < 0.001$) (see Fig. 2a and Fig. 2b). These results indicate that adolescents act

as conditional cooperators and adjust their cooperative behavior as a function of the behavioral changes of the partner. The results are in line with previous studies on adolescents [24-28]. Importantly, we found behavioral differences between the groups. In line with the hypothesis that being nice early in a relationship will result in being exploited, the main effect of shift direction was significant (F (1, 125) = 5.265, p < 0.05, η^2 = 0.04): the participants in the volatile-stable group had higher cooperation rates than those in the stable-volatile group (p < 0.05). The results suggest that the participants playing games with an initially tough partner exhibited a higher cooperation rate than those playing games with an initially nice partner. However, the interaction of shift direction × block was not significant (F (1, 125) = 3.00, p = 0.086, η^2 = 0.023). The correlation between the cooperation rates and scores on social value orientation (SVO) further confirms the construct validity of our task (r = 0.319, p < 0.001) (see Fig. 2d). We next analyzed the behavior frequencies and outcome frequencies to explore the behavioral patterns between the groups. Conducting shift direction × block × outcomes (within-subjects factor: C after C, i.e., partner cooperation following participant cooperation, vs. C after D, i.e., partner cooperation following participant defection, vs. D after C vs. D after D) ANOVAs on the behavior frequencies with Bonferroni-corrected post hoc t-tests, we found a significant interaction of shift direction × outcomes (F (3, 375) = 4.827, p < 0.01, η^2 = 0.037): compared to the participants in the stable-volatile group, the participants in the volatile-stable group cooperated more after their partner cooperated (p < 0.05) and after their partner defected (p < 0.05). In addition, the participants in this group defected less after their partner cooperated (p < 0.05) and after their partner defected (p < 0.05) (see Fig. 2e). Conducting shift direction × block × outcomes (within-subjects factor: CC, i.e., partner cooperation with participant cooperation, vs. DC, i.e., partner defection with participant cooperation, vs. CD vs. DD) three-way repeated-measures ANOVAs on the outcome frequencies with Bonferroni-corrected post hoc t-tests, we found a significant interaction of shift direction × outcomes (F (3, 375) = 4.90, p < 0.01, η^2 = 0.038): the participants in the volatile-stable group had more CC (p < 0.05) and DC (p < 0.05) outcomes and fewer CD (p < 0.05) and DD (p

< 0.05) outcomes than the stable-volatile group (see Fig. 2f). In brief, analysis of the behavior data showed that a person being nice early in the social interaction would be exploited by adolescents since our participants were generally less likely to cooperate with an initially nice partner than with an initially tough partner.

Cooperative behaviors were determined by personal intrinsic rewards derived from collaboration. Our model consisted of a learning component and a social preference component. First, a standard RL model was used to update the participants' expectations about their partners' behaviors: $p_{(t)} = p_{(t-1)}$ + learning rate × prediction $\text{error}_{(t)}$. Second, the participants made decisions by maximizing their expected value from alternative choices, which constituted not only monetary bonuses but also the social rewards derived from collaboration itself [29]; it was modeled by the following function: $U_{(\text{cooperation})}$ = expected monetary bonuses × social value. We conducted parameter recovery model prediction and result recovery to ensure model robustness and validation, and the results consistently demonstrated that our model worked well (for details, see methods, Supplementary Section 2.2 and Fig. 5). A two-sample t-test on social reward, referring to parameter ω, revealed that the volatile-stable group had a higher level of social preference for cooperation than the stable-volatile group (t (125) = -2.213, $p < 0.05$) (see Fig. 4a). In addition, increasing social reward was significantly correlated with a higher SVO score ($r = 0.28$, $p < 0.01$), which supports the construct validity of our model estimation. We also conducted the same analysis on other parameters, we found that, on average, the utility of cooperation remained significantly different between the groups (t (238) = -6.250, $p < 0.001$), but no significant effect between groups were found on other computational components (see Supplementary Section 2.2 and Fig. S5). The results suggest that the increased cooperation rate in the volatile-stable group was driven by the participants' higher intrinsic utility derived from cooperation with an initially tough person compared to an initially nice person.

Intrinsic rewards mediate the association between the shift direction and cooperative behavior. We next ran a mediation analysis to examine whether the

perceived social value from reciprocity mediates the relationship between the shift direction and cooperative behavior. The results showed that shift direction impacted the cooperation rates through changes in social value parameter ω (Sobel test, Z = 2.16, $p < 0.05$, full mediation; see Fig. 4b and Supplementary Table S3). The stepwise regression excluding shift direction was no longer significant when combined with ω (*coeff* = 0.034, t = 1.026, p = 0.307) compared to the initial coefficient (*coeff* = 0.103, t = 2.295, p = 0.023), suggesting that the shift direction effect on social value ω acted as a full mediator of the shift direction effect on cooperative behaviors. Furthermore, we ran Model 2 to confirm whether the association between SVO and the cooperation rates was mediated by social value parameter ω, and we found that the result was significant (Sobel test: Z = 3.042, $p < 0.01$, partial mediation; see Fig. 4c and Supplementary Table S4).

Discussion

Decisions to engage in cooperation can be strongly affected by the engagement of others. Studies on adolescents have consistently shown that they are more likely to cooperate when their peers are cooperative as well [24-26]. Compared to younger children, adolescents can flexibly adjust their strategy as a function of their peer's behavior; that is, they cooperate more conditionally [27,28]. Our participants consistently increased their cooperative behavior when the probability of partner cooperation was high but decreased their cooperative behavior when the probability of partner cooperation was low. These results suggest that our manipulation motivated the participants to make decisions favoring mutual cooperation. In addition, we found that adolescents scoring high in SVO exhibited a higher level of cooperation. SVO describes dispositional weights with regard to whether individuals care more for the wellbeing of another person than for their own wellbeing in a social context involving conflicts of interest [30]. Prosocial individuals exhibit greater levels of cooperation [29,31] and trust [32] toward others. Compared to proself individuals, prosocial individuals have a higher level of reciprocity, and they place more value on collective benefits and outcome equality [24]. Our results are in line with a large body of evidence showing an

association between SVO and cooperative behavior [30]. Taken together, these results support the validity and generalizability of our study.

In line with the hypothesis that being nice early in a relationship will result in being exploited, the participants interacting with an initially nice partner exhibited a significant low cooperation rate than those interacting with an initially tough partner. These results indicate the selfishness in interpersonal exchange among adolescents: imposing kindness (being initially nice) and submitting to exploitation (being initially tough). Our results are inconsistent with a previous PDG-based study from undergraduates [13], which showed that participants were less likely to cooperate with a partner exhibiting betrayal behaviors early rather than late in social interactions. This inconsistency might indicate that the social learning mechanism of adolescents is different from that of adults. Indeed, our results are in line with substantial previous studies on school bullying. Girls scoring higher in prosocial affiliation are more likely to be victimized in middle school [33]. In contrast, teenagers showing more bullying or aggressive behaviors are perceived as being more popular by their peers [34-36]. Moreover, youth with experience of maltreatment contribute more in the public goods game, which might be due to the motivation to minimize the odds of hostility and to sustain positive interactions with others at the cost of being exploited [37]. In our study, we found that the participants who encountered a tough partner were more ultimately cooperative, suggesting that they were more willing to maintain a positive relationship with a tough person.

The cooperation preferences toward an initially tough person can be attributed to alternative reasons. On the one hand, teenagers might form inappropriate expectations of their partners, such that they overestimate the material benefits derived from cooperating with an initially tough person. On the other hand, they might experience a higher level of subjective utility of reciprocity with an initially tough person compared to an initially nice person. We next employed computational modeling to dissociate these explanations by decomposing the computational subcomponents underlying cooperative behavior.

Our results showed that only the social value component was significantly different between the two groups of participants. The participants playing games with an initially tough partner received more social rewards than those playing games with an initially nice partner. These results are in line with previous findings showing that social value rather than the expectations of a partner's trustworthiness guides decision-making in trust [22] and that larger social rewards from collaboration were associated with increased activation in the ventral striatum, a key brain region for reward processing [38]. Next, the intrinsic rewards derived from collaboration fully mediated the association between the partner's behavioral patterns and the participants' responses and partially mediated the association between SVO and the participants' responses in the PDG. These results provide additional evidence that adolescents' decision to cooperate with others can mainly be attributed to the perceived social value from reciprocity. The greater intrinsic reward from collaboration with an initial-tough person might due to the self-protective motivation to avoid being exploited by others. The strategy of imposing kindness and submitting to exploitation could serve as an adaptive approach to fit with the complex social world during adolescence. In line with this idea, adolescents who do not engage in bullying events are reluctant to befriend with victims due to the fear of becoming victims themselves [39], and due to the need for help from high-status peers to maintain their own status [34]. Overall, our study demonstrated that the behaviors of peers have a great influence on adolescents' social decision-making, providing direct evidence that social rewards act as a key latent component for guiding individuals' social behavior.

In conclusion, we examined whether being nice or being tough during early social interactions would impact future social interactions. Our findings indicate that being nice is not the best strategy for adolescents to establish mutual cooperation with others and to avoid being exploited. On the contrary, being tough during early social interactions will allow one to gain more from social interactions. Computational modeling revealed that cooperation with an initially tough person is intrinsically more rewarding. This study enhances our understanding of how adolescents form a

sustained social relationship with others and may provide insights into why bullying events and conflicts are prevalent during this life stage. Our findings may also have implications for schooling and education in terms of reducing bullying, victimization and stressful experiences, such as providing more specific suggestions for designing classroom activities because we found that adolescents more cherish opportunities of cooperation after they have been exploited.

Materials and Methods

Participants

A total of 127 adolescents aged 14-17 years (mean age ± S.E. = 16.13 ± 0.06 years, 43 girls and 84 boys) were recruited from a high school in XX. The participants were randomly assigned to the stable-volatile group (n = 59) or the volatile-stable group (n = 68). All participants were screened for significant health or neurological problems and had normal or corrected-to-normal vision. Additionally, all participants and their guardians (mainly parents) provided written informed consent before the experiment. The study and participant recruitment were approved by the Ethics Committee of XX University.

Experimental procedure

On the day of the experiment, the participants were first randomly distributed to two lab rooms; each room was large enough to contain more than 70 participants. The participants were seated in the room to complete the screening and to read the task instructions. To avoid social desirability bias during the game, the cooperation task (i.e., the PDG) was always referred to as the “interaction game”. The participants were instructed that they would play multiple rounds of the interaction game in real time with another anonymous player who was putatively playing the game in the same room. They were also instructed that they and their partners would be paid in the form of gifts (such as stationery) based on their responses to several randomly selected rounds at the conclusion of the experiment. In reality, the choices of the partner were

predetermined by a computer program, such that the choices of the partner were set as cooperative in one block and as a defective in another block (see Fig. 1). Before the experiment, all participants completed several psychological measurements (e.g., the Youth Psychopathic Traits Inventory, Perceived Stress Scale and Beck Depression Inventory; for more details, see Table S1) as a routine to ensure that there were no significant differences in mental health and personality traits between the groups. Then, the participants carried out several practice rounds to ensure that they correctly understood the task. To control for the effects of mood or anxiety on cooperation, the participants also completed the Positive and Negative Affect Schedule (PANAS) [40] and State Anxiety Inventory (SAI) [41] before the game and at the end of the experiment. The participants also completed an SVO task [42] to measure the degree of prosociality of each participant depending on how selfishly or generously they weigh the wellbeing of a stranger compared to their own wellbeing.

Cooperation task

The cooperation task (**Fig. 1**) was a variant of the classical PDG with three important differences: i) to avoid social desirability bias during the game, we used geometrical shapes (i.e., a triangle or square) instead of explicit texts [43] to represent choices of cooperation or defection; ii) to capture the dynamic learning process during social interaction, the participants were led to believe that they were playing iterated games with the same person multiple rounds instead of one-shot games; and iii) to test how individuals changed their behaviors to adapt to changing environments over time, two blocks (i.e., stable or volatile cooperation pattern of the partner) as well as two different shift directions (i.e., stable-volatile or volatile-stable) were set in the current study. In the stable block, the cooperation rate of the partner remained stable (i.e., 78%) across 60 rounds. In the volatile block, the cooperation rate of the partner was varied over time, shifting between 20% and 80% every 20 rounds. The setting of the probabilities was inspired by previous studies [44]. Hence, the cooperation rate in the volatile block was more variable and unpredictable than that in the stable block. The participants were randomly assigned to two shift directions: half of the participants

experienced first the stable block and then the volatile block (i.e., stable-volatile group), whereas the other half experienced first the volatile block and then the stable block (i.e., volatile-stable group). This experiment allows us to examine our hypotheses by providing several advantages. First, the settings of two simulated partners remained exactly the same except for the direction of the shift in cooperation probability. Hence, the observed behavioral differences between the two groups of participants can be attributed to the effect driven by different shifts in behavioral patterns. Second, the experiment was conducted in large lab rooms with many participants playing games at the same time, helping to improve the ecological validity of the study. Indeed, almost all of our participants (more than 96%) firmly believed that they were interacting with a real person during the game. Finally, we leveraged recent advances in computational tools to fit our data to further understand the computational underpinnings of cooperative behavior. There were no significant differences between the two groups with respect to age and gender among other psychological dimensions (see Supplementary Table S1).

In each round, two players chose to either cooperate or defect and received a payoff, which depended on the interaction of their respective choices. Specifically, the participants were instructed to make their choice in the decision window within 4000 ms, during which they were told that their partner was making his/her choice at the same time. Then, the choices and payments of both players were presented in a feedback window for 1500 ms. Based on the choices of the two players, there were four different outcomes. Each of the four outcomes was associated with a different payoff. Specifically, participant cooperation followed by partner cooperation payed 4 tokens to both players; participant cooperation followed by partner defection payed 0 tokens to the participant and 6 tokens to the partner; participant defection followed by partner defection payed 2 tokens to both players; and participant defection followed by partner cooperation payed 6 tokens to the participant and 0 tokens to the partner. Therefore, the benefit-maximizing strategy was to always choose defection regardless of the partner's choice [45]. The participants completed 2 blocks (i.e., one stable block and one volatile block) of the iterated game, with 60 rounds in each block. After each

group of 15 rounds, three rating windows appeared, during which the participants were asked to rate their preference (i.e., *"How do you like your partner now?"*) and the cooperativeness (i.e., *"How cooperative is your partner now?"*) of their partner as well as how angry (i.e., *"How angry are you now?"*) they felt on a 9-point Likert scale, where 1 = not at all and 9 = very much.

Behavioral data analysis

The *p*-values were corrected for deviations based on Greenhouse-Geisser correction if necessary. Bonferroni correction was used to draw multiple comparisons unless otherwise noted. All statistical analyses were conducted using SPSS 21.0 (IBM, Somers, USA). We conducted sample size estimation using G*Power 3.1.9.2 [46] to determine the number of participants sufficient to detect a reliable effect. Based on an estimated average small-to-medium effect size (Cohen's d = 0.28), 103 participants were needed to detect a significant effect (α = 0.05, β = 0.80, two-by-two mixed ANOVA interaction effects). We planned to recruit 125 participants (assuming 20% of the participants would be removed from the data analysis because they might have doubts about the reality of the putative partner during the game). In the end, we recruited 127 participants. After finishing the experiment, four out of the 127 participants reported doubts over whether they were playing with a real person during the game. When probed further, these participants reported that their occasional doubt did not alter their behavior in the game. We compared the behavioral responses, reaction times, and model parameters for the participants who expressed doubt with those who did not express any doubt, and we found no significant differences. Excluding the data of these participants did not change the statistical results reported here; therefore, we included all 127 participants in our final analyses.

Cooperation rate. Two-way repeated-measures ANOVAs were conducted on the cooperation rate, with shift direction (stable-volatile vs. volatile-stable) as the between-subjects factor and block (stable vs. volatile) as the within-subjects factor.

Conditional cooperative behaviors. Based on the partner's choice in previous round

t-1 and the participant's choice in current round *t*, the outcomes can be regarded as follows: partner cooperation following participant cooperation (C after C), partner defection following participant cooperation (C after D), partner cooperation following participant defection (D after C) and partner defection following participant defection (D after D). Three-way repeated-measures ANOVAs were conducted on the behavior frequencies, with shift direction as the between-subjects factor and block and outcomes (C after C vs. C after D vs. D after C vs. D after D) as within-subjects factors.

Outcome frequencies. Based on both the partner's choice and the participant's choice in round *t*, the outcomes can be regarded as follows: partner cooperation with participant cooperation (CC), partner defection with participant cooperation (DC), partner cooperation with participant defection (CD) and partner defection with participant defection (DD). Three-way repeated-measures ANOVAs were conducted on the outcome frequencies, with shift direction as a between-subjects factor and block and outcomes (CC vs. DC vs. CD vs. DD) as within-subjects factors.

Computational modeling

Social value model with the Rescorla-Wagner (RW) learning rule. The social value model assumed that participants make decisions by i) updating their expectations about partner cooperativeness under the RL rule [47] and ii) receiving social rewards from positive interactions that facilitate future collaborations as many studies have demonstrated the rewarding nature of cooperation itself [20]. We assumed that the social reward from cooperation was conditioned on the belief regarding the partner's cooperation rate, which was updated by the choice of the partner in each round. The utility functions for cooperation and defection, respectively, are represented by the following:

$$U_{c(t)} = [a \cdot p_{(t-1)} + c \cdot (1-p_{(t-1)})] \cdot \omega \quad (1)$$

$$U_{d(t)} = d \cdot p_{(t-1)} + b \cdot (1-p_{(t-1)}) \quad (2)$$

where $U_{c(t)}$ refers to the utility of cooperation and $U_{d(t)}$ refers to the utility of defection at time *t*. Parameter ω refers to the social reward/weight on cooperation ($\omega \in [0,10]$). In addition, a, b, c and d are the monetary bonuses from the design matrix of the interaction game (i.e., a = 4, b = 2, c = 0 and d = 6). Furthermore, *p* is the inferred cooperation probability of the partner updated based on the RW learning rule:

$$p_{(t)} = p_{(t-1)} + \alpha \cdot PE_{(t)} \tag{3}$$

$$PE_{(t)} = feedback_{(t)} - p_{(t-1)} \tag{4}$$

where α is the learning rate ($\alpha \in [0,1]$), $p_{(t)}$ is the inferred probability of cooperation of the partner at time point *t* ($p \in [0,1]$), and $feedback_{(t)}$ is the choice of the partner at time point *t*, where *feedback* = 1 if the partner chooses to cooperate and *feedback* = 0 if the partner chooses to defect. $p_{(0)}$ is fixed to 0.5 for all participants. $PE_{(t)}$ is the trial-by-trial prediction error, i.e., the difference between the real choice of the partner and the participant's expectation of the probability of cooperation of the partner prior to feedback.

The probability of choosing each choice was modeled using a softmax function [48]:

$$p_{c(t)} = \frac{e^{(U_{c(t)} \cdot \beta)}}{e^{(U_{c(t)} \cdot \beta)} + e^{(U_{d(t)} \cdot \beta)}}; \tag{5}$$

where β is the softmax inverse temperature parameter, such that the lower the β is, the more diffuse and variable the choices are ($\beta \in [0,10]$). $P_{c(t)}$ refers to the probability of participants making cooperation decisions at time *t*.

The parameters for all models were estimated in MATLAB (versions of R2016b) using the fmincon optimization function for each participant by maximizing the log-likelihood of the observed data under the model on a trial-by-trial basis. We reduced the likelihood of the model converging on a local minimum by selecting 100 random start locations. The log-likelihood was calculated for each participant by

maximizing the following function:

$$\mathrm{L} = \sum_{t=1}^{n} \log(P_{c(t)} \mid (\alpha, \beta, \omega)); \quad (6)$$

where c represents the participant's choice (cooperate or defect), t represents the trial, and n is the total number of trials.

In short, three free parameters from this model were fitted via maximum likelihood estimation (MLE) based on the observed data. In this model, α refers to the learning rate from the trial-by-trial prediction error of the probability of cooperation of the partner, β refers to inverse temperature, indicating the degree of variability of the participant's decision, and ω refers to the social value from cooperation that is independent of the material bonus.

Model comparisons. We compared the social value model with other potential models by calculating the Bayesian information criterion (BIC) [49] and Akaike information criterion (AIC) [50] for each model for each participant, in addition to the protected exceedance probability (PEP) [51] for each model. We made different assumptions regarding cooperative behaviors and tested those assumptions by different models, which in turn were compared with the social value model. Several models associated with cooperation and RL were considered, including the noisy win-stay-lose-shift model [52], the choice kernel model [47], the Fehr-Schmidt inequity aversion model [4], the Roth-Erev reinforcement model and Bush-Mosteller model [53], the hierarchical Gaussian filter model [54], and a baseline model that makes assumptions regarding the participants' performance at the chance level (the details of the other models are provided in Supplementary Section 1.2). In the comparison of all 10 models, the social value model described above had the lowest BIC and AIC values, and the PEP value indicated that the social value model explained our data better than the other models. Hence, the social value model was regarded as the winning model (see Fig. 3a and Fig. 3b and Table S2), based on which the psychological parameters were estimated in subsequent analyses.

Parameter recovery. We ran parameter recovery analyses to ensure that our model

was robustly identifiable [55]. Importantly, doing so allowed us to assess whether our model was overfitting the data and to estimate the reliability of the model parameters. To that end, we created 127 simulated subjects by simulating task data at 127 random points in the space of the parameters α, β and ω and by fitting our model to these simulated subjects.

Model validation. We took additional steps to ensure that our model actually captured the data in an absolute sense with model prediction and result recovery. First, we investigated the predictions of our model by comparing the probability of cooperation of each participant based on our task (i.e., actual behavior) and that based on model prediction. To that end, first, we used the estimated parameters of the 127 participants and the experimental inputs (i.e., the behaviors of a predetermined partner) to generate individual simulated behaviors and then calculated the accuracy of the prediction by comparing the actual behaviors and the simulated behaviors. Second, we analyzed the simulated behaviors (generated by our model based on the estimated parameters of the 127 participants) to recover the effects on the behavioral observations.

Correlation analysis. Pearson's *r* correlation was analyzed to explore the possible relationship between the SVO scores and the cooperative behaviors and psychological parameters assessed by the computational models to test the validity and generalizability of our task and computational models, respectively.

Mediation analysis

Based on the results of the behavioral data and model estimation, we carried out a mediation analysis with the PROCESS macro for SPSS (Hayes, 2013) to further explore whether social value (i.e., parameter ω) mediated the influences of shift direction and SVO on cooperative behavior (i.e., cooperation rates). Specifically, we ran two mediation models: i) Model 1, in which the cooperation rates are the dependent variable (DV), shift direction is the independent variable (IV) and parameter ω is the mediator variable (MV); and ii) Model 2, which is similar to

Model 1 except that the IV is the SVO score. By convention, the indirect effect is a product of path A (the relationship of the IV with the MV) and path B (the relationship of the MV with the DV after controlling for the IV). Bootstrapping procedures confirmed the statistical significance of the specific indirect effect. Here, we generated a 95% confidence interval (CI) using 1000 samplings. If a 95% CI fails to contain zero, the indirect effect is significant ($p < 0.05$).

Statistical analysis. All statistical analyses were conducted using SPSS 21.0 (IBM, Somers, USA). For all analyses, the significance level was set to 0.05, and Greenhouse-Geisser correction nonsphericity was used when appropriate. Post hoc comparisons were evaluated using two-tailed pairwise tests with Bonferroni correction. The partial eta squared (η^2_p) [56] and Cohen's d [57] were provided to demonstrate the effect size where appropriate. The data distribution was assumed to be normal, but this assumption was not formally tested.

References

1 Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *science* **211**, 1390-1396 (1981).

2 Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J. & Sutter, M. Conditional cooperation on three continents. *Economics letters* **101**, 175-178 (2008).

3 Rustagi, D., Engel, S. & Kosfeld, M. Conditional cooperation and costly monitoring explain success in forest commons management. *science* **330**, 961-965 (2010).

4 Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *The quarterly journal of economics* **114**, 817-868 (1999).

5 Luo, Y., Eickhoff, S. B., Hétu, S. & Feng, C. Social comparison in the brain: A coordinate‐based meta‐analysis of functional brain imaging studies on the downward and upward comparisons. *Human brain mapping* **39**, 440-458 (2018).

6 Suls, J., Martin, R. & Wheeler, L. Social comparison: Why, with whom, and with what effect? *Current directions in psychological science* **11**, 159-163 (2002).

7 Dal Monte, O., Chu, C. C., Fagan, N. A. & Chang, S. W. Specialized medial prefrontal–amygdala coordination in other-regarding decision preference. *Nature Neuroscience* **23**, 565-574 (2020).

8 Chang, S. W., Gariépy, J.-F. & Platt, M. L. Neuronal reference frames for social decisions in primate frontal cortex. *Nature neuroscience* **16**, 243 (2013).

9 Craig, W. *et al.* A cross-national profile of bullying and victimization among adolescents in 40 countries. *International journal of public health* **54**, 216-224 (2009).

10 Curwen, T., McNichol, J. S. & Sharpe, G. W. The progression of bullying from elementary school to university. *International Journal of Humanities and Social Science* **1**, 47-54 (2011).

11 Fuhrmann, D., Knoll, L. J. & Blakemore, S.-J. Adolescence as a sensitive period of brain development. *Trends in cognitive sciences* **19**, 558-566 (2015).

12 Choudhury, S., Blakemore, S.-J. & Charman, T. Social cognitive development during adolescence. *Social cognitive and affective neuroscience* **1**, 165-174 (2006).

13 Lount Jr, R. B., Zhong, C.-B., Sivanathan, N. & Murnighan, J. K. Getting off on the wrong foot: The timing of a breach and the restoration of trust. *Personality and Social Psychology Bulletin* **34**, 1601-1612 (2008).

14 Anderson, N. H. & Barrios, A. A. Primacy effects in personality impression formation. *The Journal of Abnormal and Social Psychology* **63**, 346 (1961).

15 Roth, A. E. & Murnighan, J. K. Equilibrium behavior and repeated play of the prisoner\"s dilemma. *Journal of Mathematical Psychology* **17**, 189-198.

16 Skowronski, J. J. & Carlston, D. E. Negativity and extremity biases in impression formation: A review of explanations. *Psychological bulletin* **105**, 131 (1989).

17 Sherif, M. & Hovland, C. I. Social judgment: Assimilation and contrast effects in communication and attitude change. (1961).

18 Siegel, J. Z., Mathys, C., Rutledge, R. B. & Crockett, M. J. Beliefs about bad people are volatile. *Nature human behaviour* **2**, 750-756 (2018).

19 Palminteri, S., Kilford, E. J., Coricelli, G. & Blakemore, S.-J. The computational development of reinforcement learning during adolescence. *PLoS computational biology* **12** (2016).

20 Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T. & Meltzoff, A. N. The neural bases of cooperation and competition: an fMRI investigation. *Neuroimage* **23**, 744-751 (2004).

21 Fareri, D. S., Chang, L. J. & Delgado, M. R. Effects of direct social experience on trust decisions and neural reward circuitry. *Frontiers in neuroscience* **6**, 148 (2012).

22 Fareri, D. S., Chang, L. J. & Delgado, M. R. Computational substrates of social value in interpersonal collaboration. *Journal of Neuroscience* **35**, 8170-8180 (2015).

23 Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J. & Sanfey, A. G. Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive psychology* **61**, 87-105 (2010).

24 De Cremer, D. & Van Lange, P. A. Why prosocials exhibit greater cooperation than proselfs: The roles of social responsibility and reciprocity. *European Journal of personality* **15**, S5-S18 (2001).

25 Merlevede, E. *et al.* Perceptions, needs and mourning reactions of bereaved relatives confronted with a sudden unexpected death. *Resuscitation* **61**, 341-348 (2004).

26 Van Lange, P. A. & Semin‐Goossens, A. The boundaries of reciprocal cooperation. *European Journal of Social Psychology* **28**, 847-854 (1998).

27 Keil, J. *et al.* The Pizzagame: a virtual public goods game to assess cooperative behavior in children and adolescents. *Behavior research methods* **49**, 1432-1443 (2017).

28 van Hoorn, J., van Dijk, E., Meuwese, R., Rieffe, C. & Crone, E. A. Peer influence on prosocial behavior in adolescence. *Journal of Research on Adolescence* **26**, 90-100 (2016).

29 Balliet, D., Parks, C. & Joireman, J. Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations* **12**, 533-547 (2009).

30 Pletzer, J. L. *et al.* Social value orientation, expectations, and cooperation in social dilemmas: A meta‐analysis. *European Journal of Personality* **32**, 62-83 (2018).

31 Van Lange, P. A. The pursuit of joint outcomes and equality in outcomes: an integrative model of social value orientation. *Journal of personality and social psychology* **77**, 337 (1999).

32 Kanagaretnam, K., Mestelman, S., Nainar, K. & Shehata, M. The impact of social value orientation and risk attitudes on trust and reciprocity. *Journal of Economic Psychology* **30**, 368-380 (2009).

33 Nickerson, A. B. & Mele-Taylor, D. Empathetic responsiveness, group norms, and prosocial affiliations in bullying roles. *School Psychology Quarterly* **29**, 99 (2014).

34 de Bruyn, E. H., Cillessen, A. H. & Wissink, I. B. Associations of peer acceptance and perceived popularity with bullying and victimization in early adolescence. *The Journal of Early Adolescence* **30**, 543-566 (2010).

35 Parkhurst, J. T. & Hopmeyer, A. Sociometric popularity and peer-perceived popularity: Two distinct dimensions of peer status. *The Journal of Early Adolescence* **18**, 125-144 (1998).

36 Cillessen, A. H. & Mayeux, L. From censure to reinforcement: Developmental changes in the association between aggression and social status. *Child development* **75**, 147-163 (2004).

37 Keil, J. *et al.* Getting less than their fair share: Maltreated youth are hyper‐cooperative yet vulnerable to exploitation in a public goods game. *Developmental science* **22**, e12765 (2019).

38 Fliessbach, K. *et al.* Social comparison affects reward-related brain activity in the human ventral striatum. *science* **318**, 1305-1308 (2007).

39 Merten, D. E. Visibility and vulnerability: Responses to rejection by nonaggressive junior high school boys. *The Journal of Early Adolescence* **16**, 5-26 (1996).

40 Watson, D., Clark, L. A. & Tellegen, A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* **54**, 1063 (1988).

41 Spielberger, C. D. Manual for the State-Trait Anxiety Inventory STAI (form Y)(" self-evaluation questionnaire"). (1983).

42 Murphy, R. O., Ackermann, K. A. & Handgraaf, M. Measuring social value orientation. *Judgment and Decision making* **6**, 771-781 (2011).

43 Tsoi, L., Dungan, J., Waytz, A. & Young, L. Distinct neural patterns of social cognition for cooperation versus competition. *NeuroImage* **137**, 86-96 (2016).

44 Behrens, T. E., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. Associative learning of social value. *Nature* **456**, 245-249 (2008).

45 Roth, A. E. & Murnighan, J. K. Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical psychology* **17**, 189-198 (1978).

46 Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* **39**, 175-191 (2007).

47 Sutton, R. S. & Barto, A. G. *Reinforcement learning: An introduction*. (MIT press, 2018).

48 Kaelbling, L. P., Littman, M. L. & Moore, A. W. Reinforcement learning: A survey. *Journal of artificial intelligence research* **4**, 237-285 (1996).

49 Claeskens, G. & Hjort, N. L. Model selection and model averaging. (Cambridge University Press, 2008).

50 Posada, D. & Buckley, T. R. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic biology* **53**, 793-808 (2004).

51 Rigoux, L., Stephan, K. E., Friston, K. J. & Daunizeau, J. Bayesian model selection for group studies—revisited. *Neuroimage* **84**, 971-985 (2014).

52 Nowak, M. & Sigmund, K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* **364**, 56-58 (1993).

53 Ezaki, T., Horita, Y., Takezawa, M. & Masuda, N. Reinforcement learning explains conditional cooperation and its moody cousin. *PLoS computational biology* **12** (2016).

54 Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. A Bayesian foundation for individual learning under uncertainty. *Frontiers in human neuroscience* **5**, 39 (2011).

55 Wilson, R. C. & Collins, A. G. Ten simple rules for the computational modeling of behavioral data. *eLife* **8**, e49547 (2019).

56 Cohen, J. Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and psychological measurement* **33**, 107-112 (1973).

57 Rice, M. E. & Harris, G. T. Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and human behavior* **29**, 615-620 (2005).

Acknowledgments:

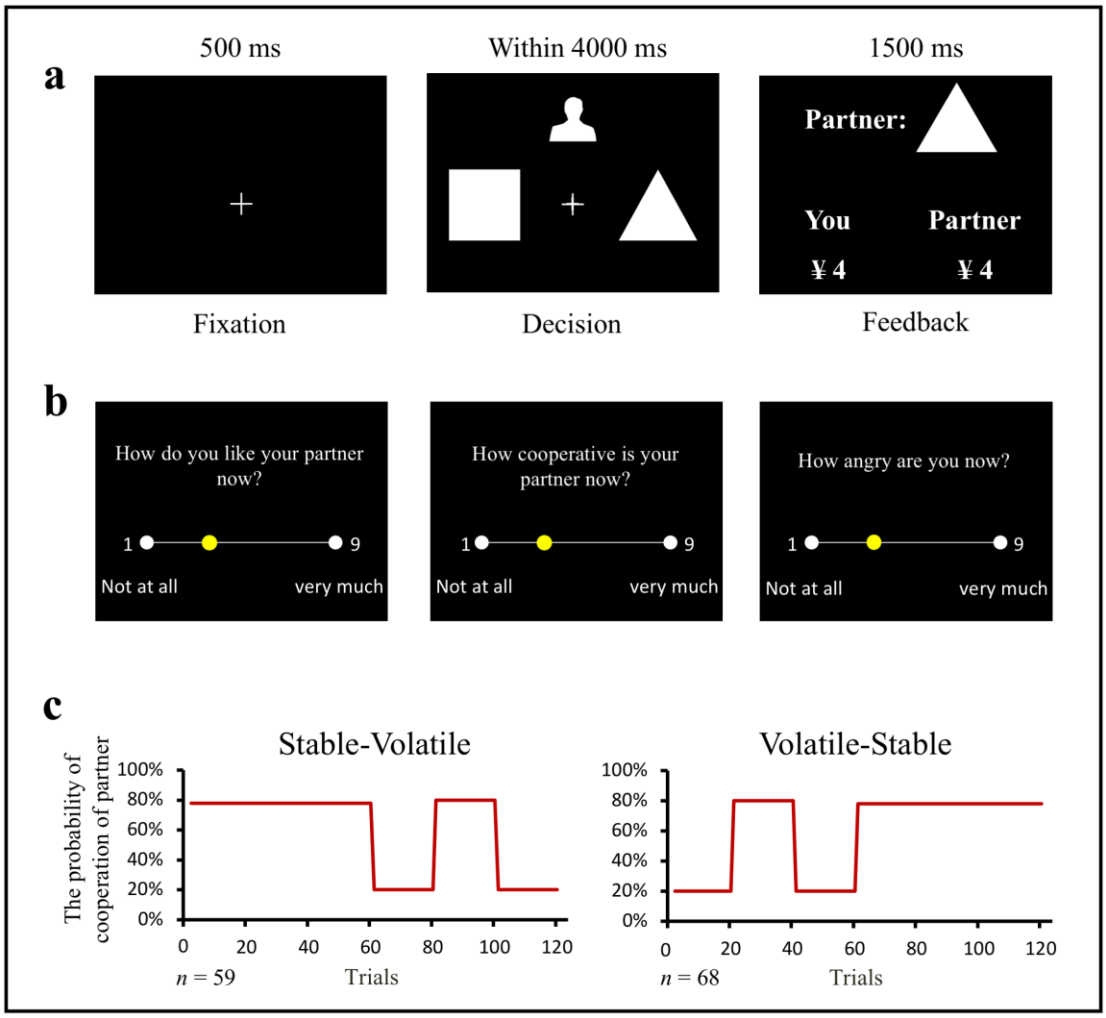

Fig. 1 | Social learning task. (a) One trial illustration. The participants make their choice by selecting one figure from two alternatives within 4000 ms. In this case, triangles and squares indicate cooperation and defection, respectively. After submitting their decision, the feedback window is presented to show the choice of their partner and the tokens received by each. **(b) Subjective ratings.** After each 15 trials, three rating windows appear. When they appear, the participants are asked to rate their preference and the cooperativeness of their partner as well as how angry they feel on a 9-point Likert scale, where 1 = not at all and 9 = very much. **(c) Task structure.** The participants are randomized to learn about how they cooperate with an agent whose behavior is stable in the first block and volatile in the second block (on the left) or with an agent whose behavior is precisely the opposite in the first and second blocks (on the right). For example, with "a stable-volatile agent", the probability of cooperation of this agent remains steady at 78% across the first 60 trials.

After the first 60 trials, the behavior of the agent becomes volatile, during which the probability of cooperation of this agent switches over each of the next 20 trials from 20% to 80%.

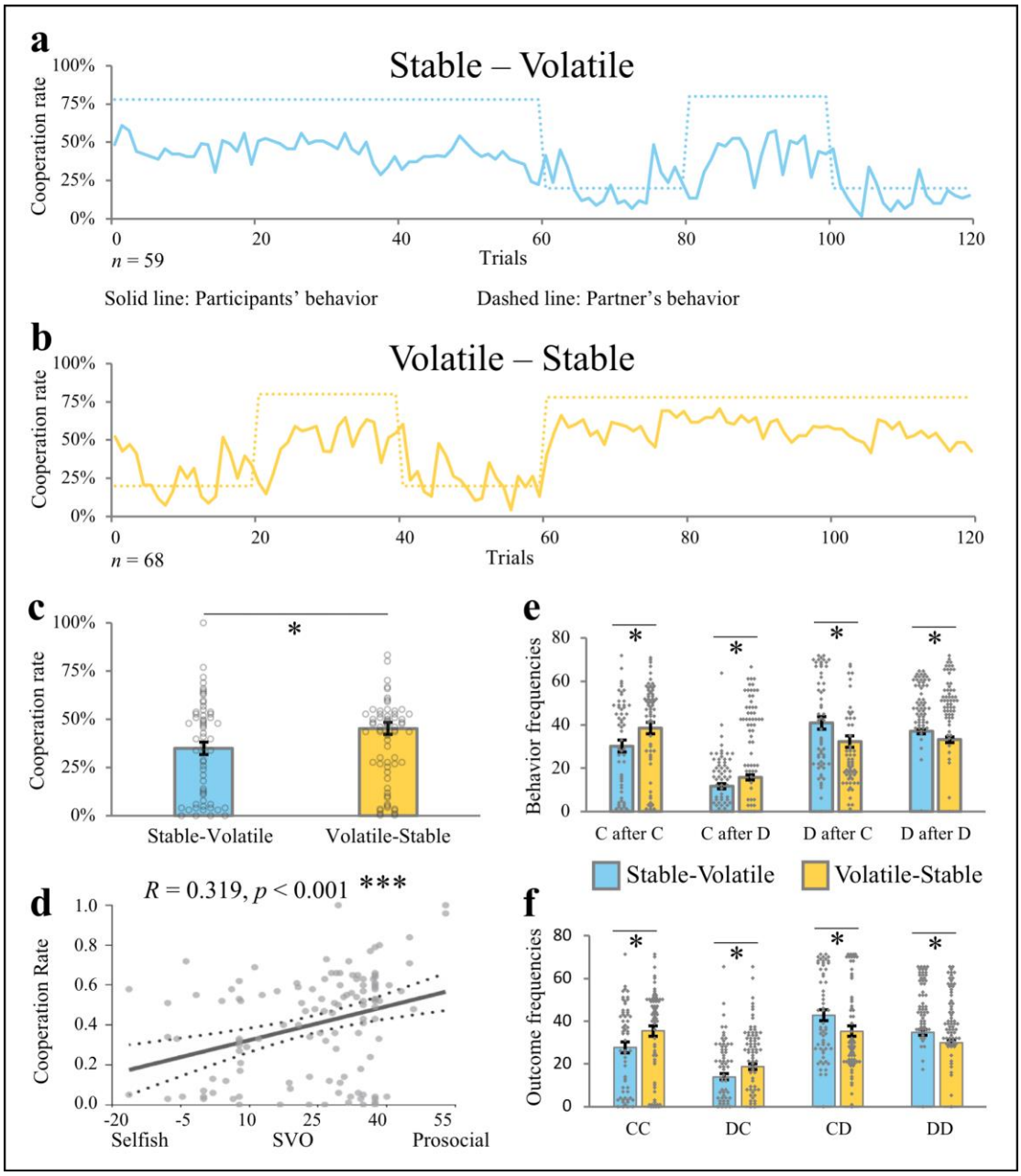

Fig. 2 | Results of the Cooperation Rate. The dotted line refers to the task structure (i.e., the partner's cooperation rate), and the solid line refers to the average cooperation rate of the participants in the stable-volatile group **(a)** and the volatile-stable group **(b)**. **(c)** The average cooperation rate of the participants in the two groups. The participants in the volatile-stable group had a higher cooperation rate than those in the stable-volatile group, especially in the stable stage. **(d)** The participants' cooperation rate was positively correlated with their SVO score. The CC, DC, CD and DD outcome frequencies for the stable-volatile group and the volatile-stable group in the stable stage **(e)** and the volatile stage **(f)**. In the stable

stage, the stable-volatile group had more CC and DC outcomes and fewer CD and DD outcomes than the volatile-stable group. Notes: ns > 0.050; *$p < 0.050$; **$p < 0.010$; ***$p < 0.001$.

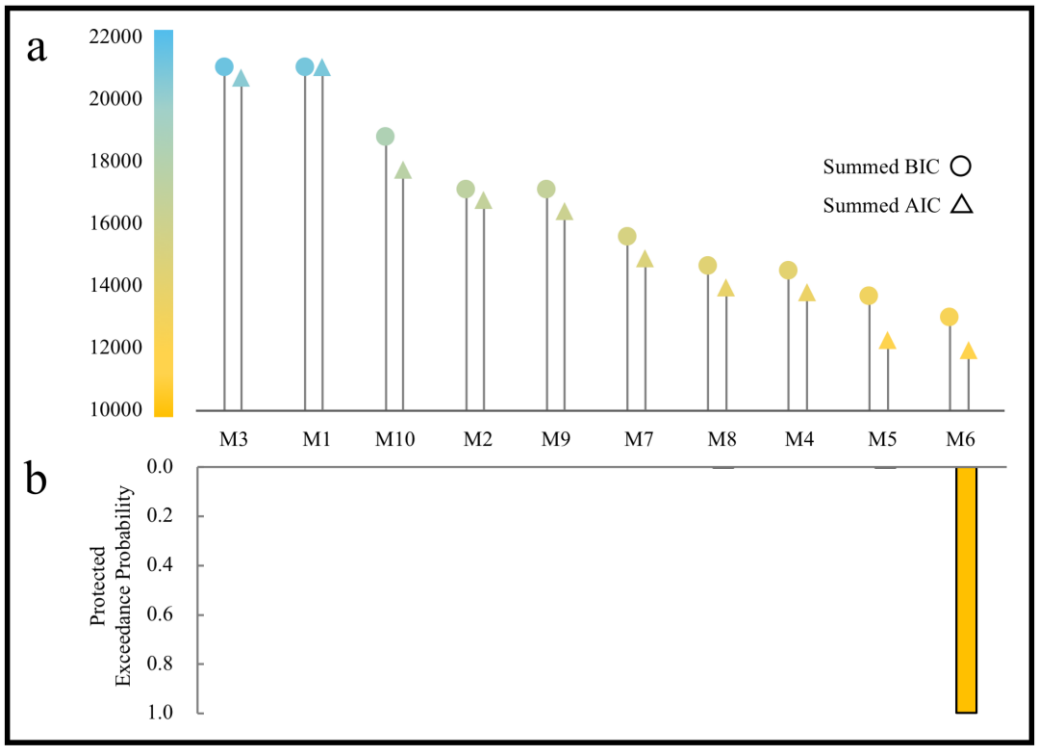

Fig. 3 | Modeling Comparisons. (a) Model comparisons based on the Bayesian information criterion (BIC) and Akaike information criterion (AIC) values. The colored dots and triangles indicate the summed BIC and AIC values for each model from the largest (left, blue) to the smallest (right, yellow). **(b)** Model comparisons based on the Bayesian model selection. The protected exceedance probabilities indicate that Model 6, a social value model with the RW leaning rule, explains our data better than the other models.

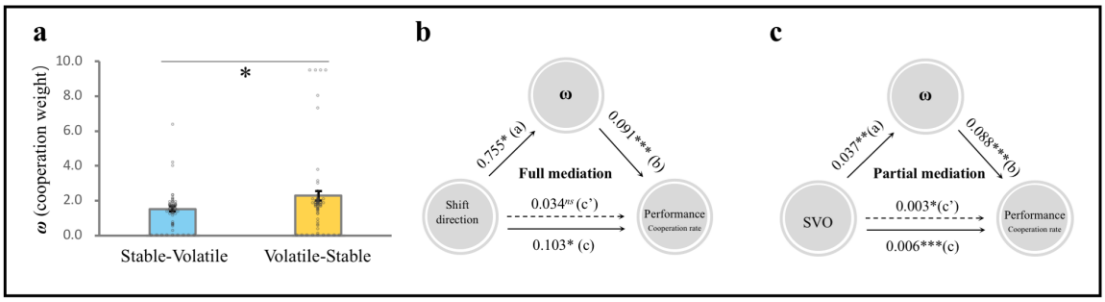

Fig. 4 | Results of model estimation. (a) The parameter estimates from the social value model with the RW learning rule. **(b)** The mediating effects of social value on the association between shift direction and social performance. **(c)** The mediating effects of social value on the association between social value orientation (SVO) and social performance. The paths are marked with standardized coefficients. Notes: ns > 0.050; *$p < 0.050$; **$p < 0.010$; ***$p < 0.001$.

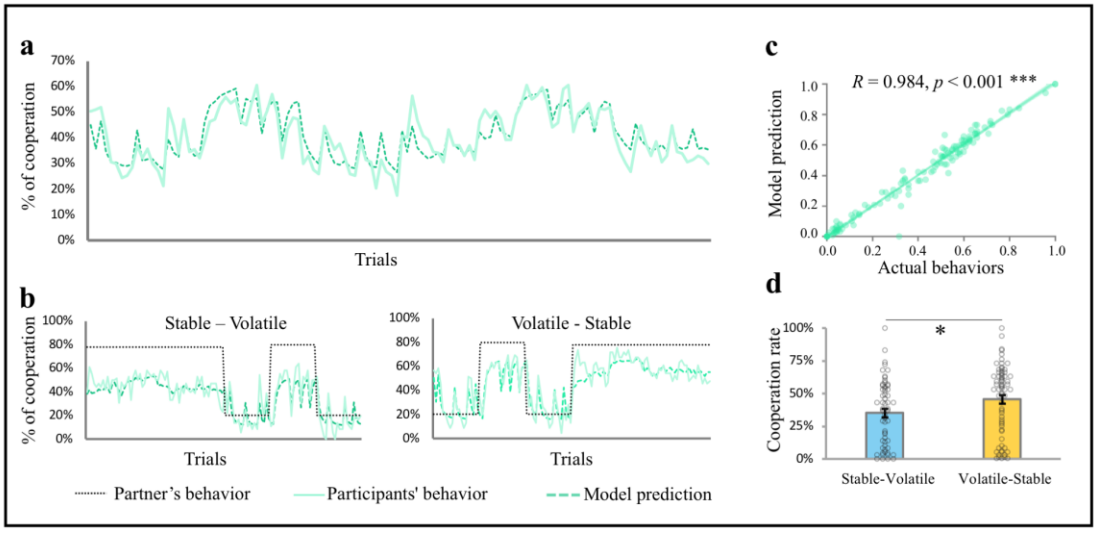

Fig. 5 | Model validation. (a) The average cooperation rate of all participants across tasks. **(b)** The average cooperation rate of the stable-volatile group (on the left) and the volatile-stable group (on the right). The green solid line refers to the actual performance of the participants, and the green dotted line refers to the simulation data generated by Model 4. **(c)** The correlation between the participant's actual behaviors and model prediction. **(d)** Result recovery from computational modeling. Notes: ns > 0.050; *$p < 0.050$; **$p < 0.010$; ***$p < 0.001$.