

# Children's restorative justice in an intergroup context

Xin Yang<sup>1</sup>  | Zhen Wu<sup>1</sup>  | Yarrow Dunham<sup>2</sup> 

<sup>1</sup>Department of Psychology, Tsinghua University, Beijing, China

<sup>2</sup>Department of Psychology, Yale University, New Haven, CT, USA

## Correspondence

Zhen Wu, Department of Psychology, Tsinghua University, Beijing 100084, China.  
Email: zhen-wu@mail.tsinghua.edu.cn

## Present address

Xin Yang, Yale University, New Haven, CT, USA

## Funding information

National Natural Science Foundation of China, Grant/Award Number: 31600899; Tsinghua University Lifelong Learning Lab; Tsinghua University Undergraduate Research Grant; Tsinghua University Initiative Scientific Research Program, Grant/Award Number: 20171080349; Tsinghua University Tang Scholar Program

## Abstract

The present study investigated 3- to 6-year-old (total  $n = 158$ ) children's restoration behaviors both when they were second-party victims (Experiment 1) and when they were third-party bystanders (Experiment 2) of transgressions. We also explored how group membership (based on color preference) affects these behaviors. We found that children preferred restoration to punishment, and that they emphasized restorative justice not only for themselves, but also for others. Furthermore, when they were victims of transgressions, the tendency to choose restoration over punishment was stronger in older than younger children. Second-party restoration behavior was influenced by group concerns in that children treated in-group transgressors more leniently than out-group and unaffiliated transgressors, but third-party restoration behavior was not. Our research challenges the view that punishment is the standard response to transgressions and suggests that alternative options like restoration are sometimes preferred over punishment by young children.

## KEYWORDS

group concern, minimal groups, punishment, restoration, restorative justice

## 1 | INTRODUCTION

Imagine someone takes something from you or from another person. How will you react? You might angrily seek revenge for yourself or for the victim, for example by taking things away from the transgressor to teach them a lesson. Alternatively, you might simply try to take your lost item back or return it to the victim without further punishing the transgressor. These actions reflect people's relative concerns between punishing the transgressor and protecting the victim. The former action focuses on severe punishment whereas the latter reflects restorative justice that focuses on undoing harm and canceling out ill-gotten gains. Despite the prevalence and potential benefits of restorative measures in our society (e.g., Saulnier & Sivasubramaniam, 2015; Weitekamp, 1993), the literature has focused on studying punishment with economic games like the ultimatum game or third-party punishment games (e.g., Fehr & Fischbacher, 2004; Jordan et al., 2014; McAuliffe et al., 2015; Wu & Gao, 2018). Very few studies have investigated people's sense of restorative justice (with few exceptions; Chavez & Bicchieri, 2013; FeldmanHall et al., 2014; Heffner & FeldmanHall, 2019; Riedl et al., 2015).

When adults are asked to choose between different response options including both restorative and punitive options, they sometimes prefer more restorative actions both when their own welfare is affected (FeldmanHall et al., 2014; especially when the transgression is not too severe, see Heffner & FeldmanHall, 2019) and when they are third-parties of the transgression (Chavez & Bicchieri, 2013; cf. FeldmanHall et al., 2014). Does such a preference develop after experiences with justice restoration, or is it an early-emerging preference prior to prolonged social learning? Testing young children before formal schooling would offer insights into this issue, as they likely have less experience compared to older children and adults. More specifically, past work looking at children's early sense of justice and fairness have shown that second-party punishment (children are victims) emerges around the age of 3 and 4 (Wu & Gao, 2018), and third-party punishment (children are bystanders) emerges between the age of 5 and 6 (e.g., McAuliffe et al., 2015). In the current study, we therefore explore how 3–6-year-old Chinese preschoolers choose between restoration and punishment. We also investigate how social relationships, that is, whether the transgressor belongs to the same or different social groups as the victim, might influence children's behavior across second-party and third-party contexts.

As far as we know, there has been only one developmental study looking at children's restoration behavior (Study 2 in Riedl et al., 2015). In their study, after an object was taken away from the victim, 3-year-olds could choose to restore the object (i.e., returning it to the victim), discard the object (i.e., throwing it away, which they call punishment), or do nothing. The critical finding was that 3-year-olds overwhelmingly chose to restore the object to the victim rather than punish the transgressor (very few children chose to do nothing). Although this finding was taken to indicate a preference for restoration over punishment, there were confounding factors that hamper strong interpretations. In Riedl et al. (2015)'s study, participants essentially chose between removing the object such that no one takes it (punishment option), and returning this object to the victim (restoration option). This design makes the punishment option similar to the restoration option (in that the transgressor always lost the object) and fairly mild (i.e., the transgressor did not lose anything owned in the first place). This might favor the restoration option: why would people throw the object away when they can return this object to the rightful owner? Therefore, it leaves open questions as to children's developing preferences for restoration versus punishment when the punishment option delivers a stronger sanction to the transgressor (thus, increasing the deterrent effect of punishment).

Furthermore, preferences for restoration or punishment may vary according to the group membership of the transgressor, which was rarely tested in prior studies. In the context of punishment research, one influential view is people generally place higher expectations on in-group members (such as the expectation to not harm but help and share resources with fellow group members; for developmental work, see e.g., Dunham et al., 2011; Rhodes & Chalik, 2013) so they might punish in-group transgressors more harshly to enforce group norms (Fehr & Fischbacher, 2004; Schmidt et al., 2012; for developmental work, see Yudkin et al., 2020). Alternatively, due to in-group favoritism, that is, preferential treatment of in-groups compared to out-groups, people might punish in-group

transgressors *less* harshly than out-groups (for a review of these two alternatives, see McAuliffe & Dunham, 2016). In-group favoritism is well-documented across various social groups (including gender, race, language, nationality, religion, and so on; e.g., Barret, 2007; Kinzler et al., 2007; Yee & Brown, 1994). It even robustly appears in lab-created, arbitrary, and meaningless 'minimal' groups (when groups are defined with meaningless terms, e.g., via a random assignment) when testing adults and children (Tajfel, 1970; for developmental work, see Abrams et al., 2008; Bigler et al., 1997; Dunham et al., 2011; Yang & Dunham, 2019; for a review, see Dunham, 2018). Some recent work suggests that this minimal in-group favoritism appears in children as young as age three (Richter et al., 2016). Therefore, our study examined whether group membership influences children's preference between restoration and punishment in 3–6-year-olds, when minimal group bias has been documented.

Which of these two accounts, in-group expectations versus in-group favoritism, fits better with people's actual behaviors? Prior studies investigating this issue have found mixed results using classic economic games like the ultimatum game, which contrasts the option of punishment with doing nothing in response to fairness violations. Generally, when people are third-parties to unfair treatments, they show in-group favoritism—they are less punitive toward in-group transgressors (e.g., Baumgartner et al., 2012; Bernhard et al., 2006; Schiller et al., 2014; but see Yudkin et al., 2020). A similar pattern was also found in 6- to 8-year-olds, suggesting that this tendency is present in children aged 6–8 (see Jordan et al., 2014), yet, we know little about how such group biased third-party punishment develops in younger children. In addition, results are less consistent when people are directly affected (as second-parties). Some studies find in-group favoritism (e.g., Kubota et al., 2013; Valenzuela & Srivastava, 2012) whereas others show the opposite pattern (e.g., McLeish & Oxoby, 2011; Mendoza et al., 2014; Wu & Gao, 2018; this is termed the 'Black Sheep Effect', see Abrams & Killen, 2014), and yet, another recent developmental study does not find group bias in either direction (McAuliffe & Dunham, 2017). These inconsistent findings may be due to different methodologies, cultures, or ages. Therefore, more studies are needed to clarify the directionality of group bias.

Overall, these above studies mainly focused on punitive response to unfairness. Meanwhile, humans also show strong tendency to restore justice, which centers on compensating the victim's welfare and canceling out ill-gotten gains (e.g., FeldmanHall et al., 2014; Heffner & FeldmanHall, 2019). Studying the priority of punishment versus restoration will greatly enrich our understanding about the principles we use when dealing with injustice, and might offer more insight into the underlying motives that drive people's behaviors in the face of unfairness. However, so far no studies have investigated how children choose between punishment and restorative options in a group context. To fill this gap, the current study directly compared children's choices both as victims (Experiment 1) and as bystanders (Experiment 2) in the face of ownership transgressions. In our study, we operationalized the punishment option such that the transgressor not only lost the object they had previously taken, but also had to pay an additional cost for the transgression. Punishment thus involved delivering a punitive sanction to the transgressor but did not provide restoration for the victim. In contrast, the restoration option involved returning the object to the victim without reducing the transgressor's original possession. Thus, this design allows us to look at children's developing preferences for restoration without *further* retribution (toward the transgressor's original possession) versus retribution without restoration.

Our primary hypothesis is that children would favor restorative options over punishment (based on past work in children, see Riedl et al., 2015, and adults, e.g., FeldmanHall et al., 2014; Heffner & FeldmanHall, 2019). This prediction is also consistent with past research on the development of children's sense of fairness (for a review, see McAuliffe et al., 2017), as children increasingly favor fair outcomes (even over having more than others) as they age. However, it is unclear whether children will choose restorative options more as victims or as bystanders, given the conflicting results in adults and little past work with children (see Chavez & Bicchieri, 2013; FeldmanHall et al., 2014). Another unexplored issue is how these preferences change with age, as past work only examined adults' behaviors and 3-year-olds' behaviors in separate studies (Chavez & Bicchieri, 2013; FeldmanHall et al., 2014; Riedl et al., 2015). Thirdly, the current study explored how the minimal group membership of the transgressor affects children's behaviors. Given the inconsistent findings on group norm enforcement (leading

to harsher punishment on in-groups) and in-group favoritism (leading to less harsh punishment on in-groups), it remains unclear what patterns of group differences would emerge in our studies. Also, as minimal group studies are mostly conducted in Western, individualistic societies, it is important to extend this work to different cultures; in particular, in adults, collectivist cultures may show less in-group bias in the minimal group setting (Falk et al., 2014). Overall, the present study aimed to fill in the gap by investigating how Chinese preschoolers choose between restoration and punishment when facing norm violation in a group context, and how such preferences differ between behaving as second-party victims and third-party bystanders.

## 2 | EXPERIMENT 1

### 2.1 | Methods

#### 2.1.1 | Participants

Power analysis using G\*Power 3.1 showed that we needed  $n = 90$  to be able to have >80% power to detect the main effect of children's preference between restoration and punishment (two tails,  $\alpha = .05$ , assuming small to medium effect size  $d = .3$ ). We recruited 94 participants in order to account for potential exclusions. Participants were tested in a quiet room either in the lab or at local preschools in Beijing, China. We compared 3- to 4-year-olds and 5- to 6-year-olds, because past work finds that the latter age group (but not the former one) showed group biases in second-party punishment (Wu & Gao, 2018). Also, the ability to infer others' mental states (which plays an important role in children's fairness-related behaviors, e.g., Takagishi et al., 2010; Tsoi & McAuliffe, 2019; Wu & Su, 2014) undergoes critical developmental changes between the age of 4 and 5 (Wellman et al., 2006; Wimmer & Perner, 1983). The final sample included 86 preschool children, with 43 3- to 4-year-olds ( $M = 4.30$  years,  $SD = 0.48$ , range = 2.96–4.95 years, 24 female) and 43 5- to 6-year-olds ( $M = 5.58$  years,  $SD = 0.47$ , range = 5.00–6.55 years, 21 female). An additional eight children were tested but excluded from data analyses for experimenter error ( $n = 4$ ), refusal to complete the experiment ( $n = 3$ ), or failure to pass the comprehension questions ( $n = 1$ ). Experiments reported in this article were approved by Tsinghua University Institutional Review Board (protocol #201602). Parental consent was obtained via the preschools in advance of all testing.

#### 2.1.2 | Design

We employed an experimental set-up that is at least somewhat more naturalistic and understandable to children as young as three, particularly in that it does not involve as much currency/market-based reasoning as economic games. This is an important consideration given that recent findings suggest an earlier emergence of certain forms of punishment behaviors (i.e., costly third-party punishment) in more naturalistic settings (Yudkin et al., 2020). There were two experimenters in the testing room: E1 was the main experimenter and E2 only operated puppets. Children were tested individually. They played three one-on-one games with three puppets in a counterbalanced order (Group, within-subject): one puppet was the child's in-group member, one was an out-group member, and one had not joined either group (i.e., hereafter 'unaffiliated' partner; following adult work that introduces a group-neutral member, Schiller et al., 2014).

There were two types of trials: At the beginning of each trial, the child and the puppet were each given two tokens. In the partner-intervening trials, the puppet stole a token from the child without permission; in the experimenter-intervening trials E1 moved one token from the child to the puppet (but the child participants could not direct any action toward E1 in the game setting). Therefore, both types of trials resulted in one token for the child and three tokens for the puppet, creating a similar unfair outcome for the child; however, the

critical difference was that the partner violated ownership rule in the partner-intervening trials but not in the experimenter-intervening trials. We designed the experimenter-intervening trials as the control situation: children's behaviors there demonstrated how frequently they wanted to try different response options (as discussed in McAuliffe et al., 2015) and how they preferred different distributions in the absence of antisocial behavior by the puppet. Therefore, the difference between the two trial types would more clearly reveal children's punitive and restorative motives. There were three trials of each type, resulting in six trials with each puppet. Trials were presented in a randomized order.

The child was given two options: restoring equality by taking the token back (leading to two tokens for each, the *restoration* option) or costly punishing the puppet such that the puppet lost all three tokens while the child lost one token (leading to the final outcome of one token left for the child but zero for the puppet; the *punishment* option). These options were made by pushing one of the two cars placed on the apparatus, which enacted the choice in a clear, salient and fun manner (Figure 1). Children's choices were dichotomously coded (0 = punishment, 1 = restoration) for data analyses.

### 2.1.3 | Apparatus

See Figure 1.

### 2.1.4 | Procedure

#### *Group assignment*

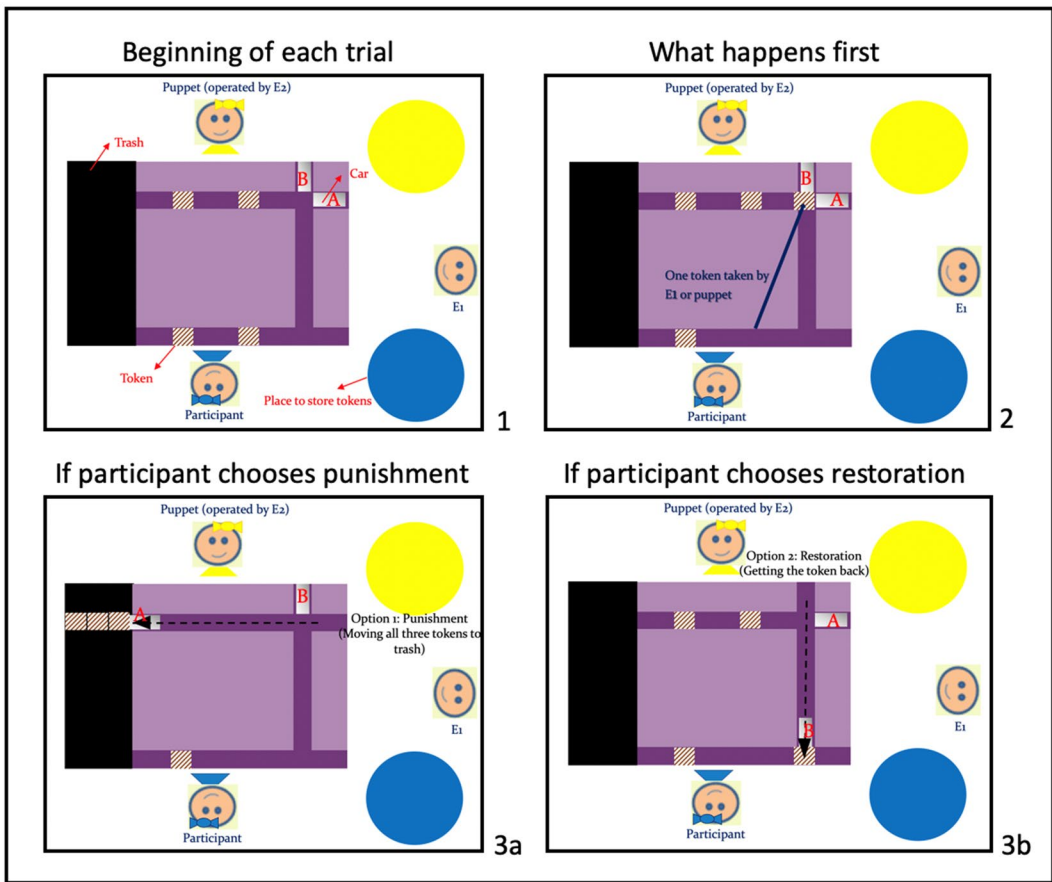
Before the experiment, the child joined the 'blue' or 'yellow' team based on his or her self-reported color preference. The child picked the color s/he preferred from two pairs of little wooden shoes on the table (one pair was blue and the other pair was yellow), and then, s/he was given the shoe s/he chose and was assigned to that color group (52% participants chose blue). Next, the child put on a team shirt, a team waistband, and a team mascot (gender-matched headwear: girls wore bowknots and boys wore crowns), and s/he also received a storage board (all items were color-matched). Using multiple visual group markers was aimed to elicit an in-group bond in young children (Richter et al., 2016). We used the word 'team' with children in order to help them better understand group memberships and to better induce group bias.

#### *Group introduction*

E1 told the child that s/he would play one-on-one games with three 'kids'—which were actually puppets operated by E2—one on the blue team, one on the yellow team, and one that was unaffiliated with either team (we always referred to the puppets as 'kids' with the child; the unaffiliated puppet was described as someone who had not decided about which team to join). The blue-team and yellow-team puppets also had the color-matched group markers (i.e., toy shoes, shirts, waistbands, mascots, and the storage boards for blocks), but the unaffiliated puppet only had the storage board and importantly, it matched the apparatus' color (i.e., purple) in order not to give it any sort of group identity or involve any new color. In this phase, the child was familiarized with the three puppets. Puppet assignment (i.e., which puppet was assigned to which group) and sequence of group introduction were counterbalanced across participants.

#### *Game introduction*

The game rule and the novel apparatus were introduced to both the child and the puppets by E1. At the start of each trial, both the child and the puppet had two tokens, which they knew they could exchange for stickers in a one-to-one fashion after the game. To operationalize the randomized trial order in an engaging manner, E1 asked



**FIGURE 1** The layout of the apparatus. In this example, the participant was in the blue group, playing the game with an outgroup puppet. Each trial began with E1 giving both sides 2 tokens (i.e., wooden blocks). Then either E1 (in the experimenter-intervening trials) or the puppet (in the partner-intervening trials) took a token away from the participant, resulting in 1 token for the Participant and 3 tokens for the puppet. If the child moved Car A on the right side of the apparatus (Option 1: punishment), it resulted in removing all 3 tokens to the inaccessible black area (final outcome: Participant had 1 and Puppet had 0). If the child moved Car B on the far side of the apparatus (Option 2: restoration), it resulted in restoring the lost token back to the child (final outcome: Participant had 2 and Puppet had 2)

the child to pick one stick from six wooden sticks without replacement, which determined the order of trials. To demonstrate the outcome of two trial types, E1 explained that if the child picked a marked stick, then, she would move one of the child's tokens to the puppet; by contrast, if the child picked an unmarked stick, then, E1 said that she would not take any action, but the puppet would move one of the child's tokens without permission. Whichever case happened, the distribution of tokens changed from 2:2 (two tokens for each one) to 1:3 (one for the child and three for the puppet), creating disadvantageous inequality for the child.

The two trial types were demonstrated to children in a counterbalanced order, and then, E1 asked a set of comprehension questions (about the number of tokens himself/herself and the other puppet would receive if s/he chose punishment or restoration options; 'You can choose to push this car or that car. If you push this car [demonstration], how many tokens will you receive? How many tokens will your partner receive? If you push that car [demonstration], how many tokens will you receive? How many tokens will your partner receive?'). E1 always

asked children how many tokens they would receive before asking how many tokens the puppet would receive. Children had to pass all the comprehension questions before the *Experimental Session*. Only one participant failed the comprehension check and was excluded.

The words 'punishment/punish', 'restoration/restore', or 'thief/steal' were never mentioned during the game; instead, E1 directly demonstrated the options and outcomes with cars and tokens.

#### *Experimental phase*

The child played three one-on-one games following the same order as in *Group Introduction*. In each game, s/he was asked by E1 to greet the puppet (operated by E2) and was reminded of the puppet's group membership. The s/he played the game as described above. More specifically, when the participant picked a stick with a mark, E1 said 'This stick has a mark on it, so I will move one of your tokens' and then, performed the action. When the participant picked a stick without a mark, E1 said 'This stick does not have a mark on it, so I will not move your tokens'. Next, the puppet partner said 'I will take one of your tokens' while performing this action. After the tokens were moved, E1 asked the participant 'Which car do you want to push?' and recorded behavioral response on a piece of paper. After each game, the child said goodbye to the puppet, exchanged for the stickers, and then, started the game with the next puppet.

#### *Group bias measures in manipulation check*

After playing with all three puppets, E1 showed the child all puppets again and asked the following questions and explanations for his/her answers: (a) *Who do you like the best?* [Explicit attitude]; (b) *Here is a sticker and you can give it to one of them. Who do you want to give it to?* [Resource allocation]; (c) *One of these three kids helped her/his mom do chores today. She/he is a good kid. Who did the chores?* [Behavioral attribution].

These three questions were used to check whether we successfully induced intergroup bias, as past research has shown that children of similar ages generally express explicit liking for in-groups, share more with in-groups, and attribute more positive behaviors to in-groups (Dunham et al., 2011; Yang & Dunham, 2019).

#### *Final explanation*

In the end, E1 asked the child why she pushed certain car(s) in the last trial and recorded their answers. We treated this as an exploratory measure and only reported descriptive results in Supporting Information.

### 2.1.5 | Coding and reliability

E1 videotaped the experiment and manually coded the children's behavioral data (punishment = 0, restoration = 1) and answers to manipulation check questions during the experiment. A research assistant, who was unaware of the study design and hypothesis, independently coded 30% of the videos. Interrater agreement was 100% (for behavioral data, Cohen's  $\kappa = 1.00$ ; for manipulation check questions, Cohen's  $\kappa = 1.00$ ).

## 2.2 | Results

We first report results on manipulation check questions, then, main analyses examining children's choices between restoration and punishment options and the developmental trajectory. Preliminary analyses revealed no effect of participant gender, color group assignment, or order (all  $ps > .27$ ), so these factors were not included in our models. Exploratory analyses of trial number and open-ended explanations are reported in Supporting Information.

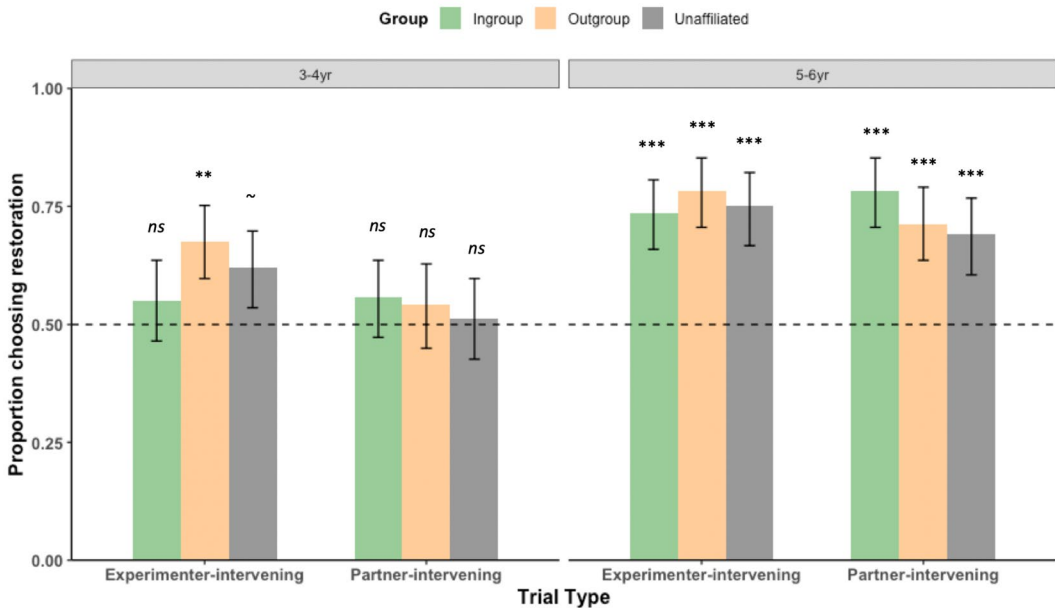
## 2.2.1 | Group bias in manipulation check

Children chose the in-group puppet 47% of the time, the out-group puppet 28% of the time, and the unaffiliated puppet 25% of the time. A chi-square test showed that they chose the in-group puppet significantly more frequently than the other two options,  $\chi^2(2, N = 86) = 22.07, p < .001$ , indicating that our manipulation succeeded in inducing in-group bias.

## 2.2.2 | Restoration versus punishment choices

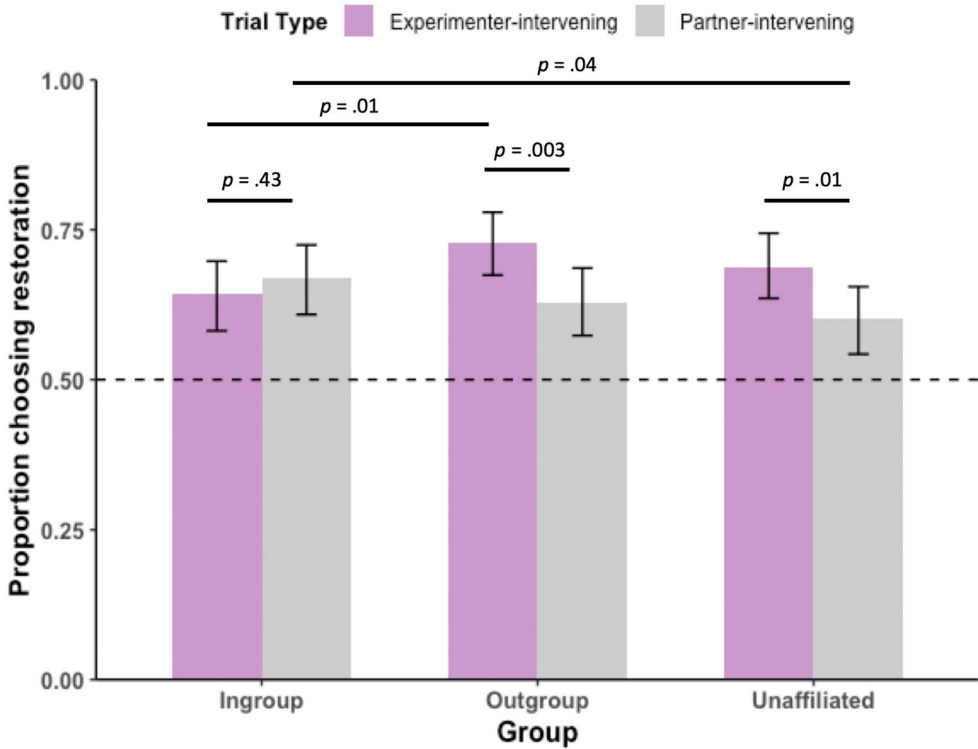
We fit a binomial mixed effects model predicting children's choices (0 = punishment, 1 = restoration) as a function of group (in-group, out-group, and unaffiliated), trial type (partner-intervening and experimenter-intervening), age group (3–4 years and 5–6 years), and their interactions, with a random intercept for participant. We found a significant group by trial type interaction (likelihood ratio test,  $\chi^2(2, N = 86) = 8.16, p = .02$ ) and an effect of age group ( $\beta = 1.25, SE = .42, 95\% \text{ CI } [.43, 2.07]$ ). The effect of group was not significant (likelihood ratio test,  $\chi^2(2, N = 86) = 2.05, p = .36$ ). Central to our primary focus, older children were more likely to prefer restoration over punishment. As shown in Figure 2, children by the age of 5 and 6 clearly preferred restoration over punishment (all  $ps < .001$ , intercepts comparing to chance) whereas 3–4-year-olds did not show a clear preference. Analyses treating age as a continuous variable also showed similar results ( $\beta = .90, SE = .26, 95\% \text{ CI } [.38, 1.42]$ ).

The group by trial type interaction was driven by significant effects of trial type with out-group and unaffiliated puppets but not with the in-group puppet. As shown more clearly in Figure 3, compared to the experimenter-intervening trials, children directed restoration at out-group and unaffiliated puppets less (directed punishment



**FIGURE 2** Experiment 1: Forced-choice between punishment and restoration across trial types (experimenter-intervening and partner-intervening trials), groups (ingroup, outgroup, and unaffiliated partners), and age groups (3–4 years and 5–6 years). Y-axis shows the proportion choosing restoration instead of punishment, ranging from 0 to 1. Error bars are bootstrapped 95% confidence intervals





**FIGURE 3** Experiment 1: Interaction between group and trial type averaged across participants (two age groups were combined because there was no interaction involving age group). Y-axis shows the proportion choosing restoration instead of punishment, ranging from 0 to 1. Error bars are bootstrapped 95% confidence intervals

more) in the partner-intervening trials (out-group,  $\beta = -.66$ ,  $SE = .22$ , 95% CI [-1.10, -.22]; unaffiliated,  $\beta = -.54$ ,  $SE = .22$ , 95% CI [-.97, -.11]), suggesting stronger punitive motives toward out-group and unaffiliated partners in partner-intervening trials. However, this punitive motive weakened for the in-group puppet, as children chose restoration and punishment similarly with the in-group across the two trial types ( $\beta = .17$ ,  $SE = .22$ , 95% CI [-.26, .60]). Unexpectedly, in experimenter-intervening trials children chose restoration less (punishment more) with in-group partners than with out-group ones,  $\beta = -.57$ ,  $SE = .23$ ,  $p = .01$  (the other pairwise comparisons were not significant,  $ps > .19$ ).

Adding children's in-group bias levels shown in manipulation check questions (answering in favor of the in-group puppet on zero to three questions) to the model did not reveal an interaction between bias and group ( $p = .92$ ). Children who showed different levels of in-group bias on the manipulation check did not respond differently toward different groups in choosing restoration and punishment. To demonstrate this finding in a different way, we also classified children into three categories based on how they chose between punishment and restoration with in-group and out-group members: those who punished the in-group less than the out-group ( $N = 19$ ), those who punished the in-group more than the out-group ( $N = 21$ ), and those who punished the in-group and the out-group at the same degree ( $N = 46$ ). The three categories of children did not differ in their levels of in-group bias on the manipulation check (one-way ANOVA,  $p = .13$ ). Taken together, children's in-group biases on the manipulation check questions concerning in-group preference did not seem to be related to their responses toward different groups during the game.

## 2.3 | Discussion

We successfully induced in-group biases in 3- to 6-year-old Chinese preschoolers and measured their punishment and restoration behavior. Results showed that children chose restoration (i.e., taking their object back/the '2:2' option) more frequently than punishment (i.e., punishing the partner/the '1:0' option). We also found that 5- to 6-year-olds chose restoration more than did 3- to 4-year-olds, a developmental pattern first shown in the field. Furthermore, we showed that children's restoration and punishment behaviors were affected by group membership (via an interaction with trial types): when the puppet intervened and created an unfair outcome compared to when it did not, children were more likely to punish the out-group and unaffiliated puppets (i.e., they chose restoration less), but not the in-group puppet. The direction of this group effect suggests in-group love but not out-group hate; that is, the difference between the in-group and the out-group was driven by the in-group moving away from the neutral baseline, not the out-group moving away from the neutral baseline.

Importantly, this is one of the first studies that demonstrate a preference for restoration over punishment in young children and the first study that reveals an age-related increase in that preference (see also Riedl et al., 2015). Children might develop an early preference for restoration because they view it as a better alternative to harsh punishment, that is, it restores justice by delivering a moderate level of cost to the violator while also compensating the victim. Our findings also imply that, when restoration is available (as in our paradigm), children increasingly favor restoration with age. Our study suggests that it may be important to provide participants with both punishment (violation-oriented) and restoration (victim-oriented) options in order to reveal a more complete picture (see also Chavez & Bicchieri, 2013; FeldmanHall et al., 2014; Heffner & FeldmanHall, 2019).

This is also one of the first studies to demonstrate something much like a classic minimal group effect in non-Western children (see also Wu & Gao, 2018), and suggests that the tendency to immediately prefer social in-groups is culturally widespread.

## 3 | EXPERIMENT 2

Experiment 1 found that when children's own interest was at stake, children aged 5–6, but not 3–4, robustly chose restoration over punishment. Experiment 2 aimed to replicate and extend the findings in the second-party context (children as victims) to a third-party context (children making choices for a victim). In a third-party context, children's choices may more clearly reveal the balance between restorative and punitive motives without being affected by their self-interested concerns. As intervening as a third-party appears later in development (e.g., between the age of 5 and 6 in third-party punishment; e.g., McAuliffe et al., 2015), here, we focused on 5-year-olds and 6-year-olds.

### 3.1 | Method

#### 3.1.1 | Participants

Power analysis using G\*Power 3.1 showed that we needed  $n = 71$  to be able to have > 80% power to detect the main effect of children's preference for restoration over punishment (because we had prior prediction of direction based on Study 1 and Riedl et al., 2015;  $\alpha = .05$ , expecting small to medium effect size  $d = .3$ ). We recruited 81 participants in order to account for potential exclusions. Participants were tested in a quiet place at a local preschool in Beijing, China. The final sample included 72 preschool children ( $M = 5.99$  years,  $SD = 0.38$ , range = 5.13–6.68 years, 36 females). There were 36 5-year-olds ( $M = 5.66$  years,  $SD = 0.23$ , range = 5.13–5.99 years, 19 females) and 36 6-year-olds ( $M = 6.32$  years,  $SD = 0.19$ , range = 6.03–6.68 years, 17 females) by median split. An additional

nine children were tested but excluded from data analyses for experimenter error ( $n = 3$ ), failure to complete the experiment ( $n = 3$ ), or failure to pass the comprehension questions ( $n = 3$ ). Parental consent was obtained via the preschool in advance of all testing.

### 3.1.2 | Design

For simplicity of design, the victim puppet was always the child's in-group member whereas the other puppet partner was from an in-group, an out-group, or neither group (unaffiliated). The child was asked to help the in-group victim puppet choose what to do. The child always received a fixed amount of payoff for participation regardless of the choices they made. Other changes from Experiment 1 were detailed in Procedure.

### 3.1.3 | Apparatus

Similar to Figure 1 except for minor procedural changes that are detailed in the following section.

### 3.1.4 | Procedure

There was a slight change of order, in that we added a manipulation check before all testing to probe baseline intergroup attitudes, and that we introduced the game rule before introducing the puppets (there were four puppets in this experiment). We made the partner-intervening trials more salient by highlighting the fact that the partner stole the possession of the victim against its will (see below). We used a blocked design for the two trial types rather than a trial-by-trial lot-drawing procedure to avoid confusion. We also made a few other changes to clean up procedures and avoid potential confounds, detailed below.

#### *Group assignment*

The child was asked about his/her preferred color between 'blue' and 'yellow' (51% participants chose blue). Instead of a storage board, the child received a non-transparent storage box with a color-matched lid so that s/he could not easily see how many blocks each one has, decreasing the possibility of comparing the number of tokens with each other.

#### *Group bias measures in manipulation check*

Instantly after *Group Assignment*, the child was shown six identical drawings of two side-by-side kids (one wearing blue markers and the other one wearing yellow markers) sequentially. We first checked that s/he knew which groups these kids belonged to and whether they were his/her in-groups. After s/he had successfully answered this experimenter-intervening question, we asked three 'Explicit attitude' questions ('*Who do you like better?*') as we showed s/he the first three drawings one by one. Then, we asked three 'Behavior attribution' questions as we showed s/he the last three drawings (the order of the three questions were counter-balanced across participants): (a) '*One of these kids helped her mum with housework at home today. Who did the housework?*' (b) '*One of these kids helped another kid at the preschool today. Who helped that kid?*' (c) '*One of these kids brought juice for her friends today. Who brought juice for her/his friends?*' These three questions were modeled after past research (Dunham et al., 2011), which were shown to measure in-group biases in young children (note that we included more questions than in Study 1 to better measure in-group bias). Following prior work (Gonzalez et al., 2020; Jordan et al., 2014; Wu & Gao, 2018), we added these questions before the main task to make sure the minimal group manipulation was successful before administering main behavioral measures.

### *Game introduction*

We introduced the child to the game which the other kids would come to play with s/he later. For ease of demonstration, E allocated the tokens unevenly, putting one token on one side and three tokens on the other side. The child was told that the tokens put on each side were given to the kid close by during the game (but during the demonstration there was no other 'kid'), and one token could be exchanged for one sticker. We also told the child that s/he would receive one page of stickers at the end of the game regardless of the outcome of the game (all participants said that they loved stickers, and we told them that the puppets loved stickers too). Comprehension questions were similar to those of Experiment 1 except for some minor changes in wording to refer to the two absent 'kids'. All but one participant passed the comprehension questions (who was then excluded from data analyses).

### *Puppets introduction*

The in-group victim was always introduced first to the child; the order of introducing the other three puppets acting as the in-group, out-group and unaffiliated violators was counter-balanced across participants.

### *Experimental phase*

The in-group victim, standing closest to the child, played one-on-one games with three puppets following the same order as in *Puppet Introduction*. In the partner-intervening trials, E gave each of the puppets two tokens. Then, the in-group victim left because it needed to go to the restroom, eat some food, or drink some water (the order of excuses was counterbalanced). After the in-group victim left, the puppet partner took one of its tokens away and put it on its own side. Next, the victim returned and found that its token was taken away. E asked the child to help it push a car ('He/She came back and found that one of his/her token was missing. But he/she couldn't reach the cars. Could you help him/her push a car?'). This change aimed to make it clear that the partner stole a token from the in-group victim against its will. In the experimenter-intervening trials, to match the outcome distribution, E gave one token to the in-group victim and three tokens to the other puppet at the beginning of the trial. Unlike Experiment 1, here, E did not change a 2:2 distribution to 1:3 by moving one of the tokens, so that we made the experimenter-intervening trials more distinctive in actions from the immoral trials but with the same outcome. Again, children could only punish the puppets but not the experimenter. Then, E asked the child to push a car ('Which car do you want to push?').

### *Explanation, memory check, evaluation, and final manipulation check*

In the end, E1 asked the child the following questions:

1. Why s/he pushed certain car(s) in the games [*Explanation*]? As in Experiment 1, we treated this as an exploratory measure and only reported descriptive results in Supporting Information.
2. What the three puppets [point to the three partners] did in the games [*Memory Check*]? All participants were able to answer this question.
3. Evaluations of the partners (was good or bad) and emotion judgments of the victim (felt happy or sad). Each answer was followed with a question to measure the scale (e.g., was it a little bad or really bad?). We coded their answers on a 4-point Likert-like Scale [very bad = 1, very good = 4; very sad = 1, very happy = 4]. We always pointed to the certain puppets and called its name when we asked questions about it. Pictures of 'thumb-up' and 'thumb-down' (for evaluations), or 'smiling face' and 'crying face' (for emotions) were used to help children answer these questions [*Evaluation*]. We added this as an exploratory measure and report analyses of these questions in Supporting Information.
4. (Among the three partners) *Who do you like the most?* [*Final Manipulation Check*].

### 3.1.5 | Coding and reliability

E videotaped the experiment and manually coded children's behavioral data (0 = punishment, 1 = restoration) and answers to manipulation check questions during the experiment. A research assistant, who was unaware of the study design and hypothesis, independently coded 30% of the videos. Interrater agreement was excellent (for behavioral data, *Cohen's*  $\kappa = 1.00$ ; for manipulation check questions, *Cohen's*  $\kappa = 1.00$ ).

## 3.2 | Results and discussion

Data were analyzed in the similar models as specified in Experiment 1. Preliminary analyses revealed no effects of trial number, participant gender, or color group assignment, so we did not include these factors in the models (however, there were some unexpected order effects; they were detailed in Supporting Information). We also reported analyses of open-ended explanations and evaluation and emotion rating questions in Supporting Information.

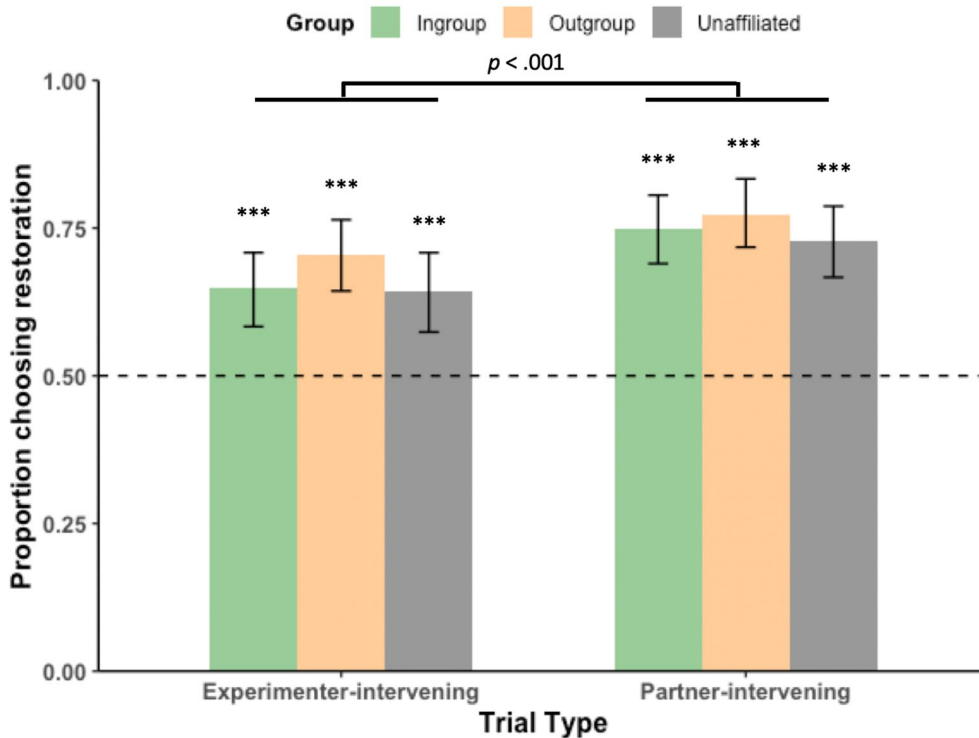
### 3.2.1 | Group bias in manipulation check

We first analyzed the manipulation check questions to confirm that intergroup bias was successfully induced. In the pre-game check (6 questions), children chose the in-group member on 64.58% of the questions (forced-choice between an in-group and an out-group member; chance = .5),  $t(71) = 7.18$ ,  $p < .001$ , *Cohen's*  $d = .85$ . In the post-game check (choosing among 3 puppets), 47 children said they liked the in-group puppet the most, although only 17 favored the out-group and 7 chose the unaffiliated,  $\chi^2(2, N = 72) = 36.62$ ,  $p < .001$ . These results showed that our manipulation succeeded in inducing in-group bias.

### 3.2.2 | Primary results on restoration versus punishment choices

As in Experiment 1, we fit a binomial mixed effects model predicting children's choices (0 = punishment, 1 = restoration) as a function of group (in-group, out-group, and unaffiliated), type (partner-intervening and experimenter-intervening), age group (5 and 6 years), and their interactions, with a random intercept for participant. There was no effect of age group ( $\chi^2(1, N = 72) = .64$ ,  $p = .42$ ), nor was there an effect of age if we included age as a continuous variable ( $\chi^2(1, N = 72) = .93$ ,  $p = .34$ ), so we collapsed both age groups.

As shown in Figure 4, the above model showed that children preferred restoration over punishment ( $ps < .01$ , intercepts comparing to chance). We also found a significant effect of trial type,  $\beta = .63$ ,  $SE = .15$ , 95% CI [.34, .92]: children directed more restoration in the partner-intervening trials compared to the experimenter-intervening trials. In addition, the effect of group was not significant ( $\chi^2(2, N = 72) = 4.94$ ,  $p = .08$ ). Adding children's in-group bias levels shown in manipulation check questions (answering in favor of the in-group puppet on 0 to 6 questions on pre-game manipulation checks) to the model again revealed that the interaction between bias and group membership was not significant ( $p = .28$ ). As in Experiment 1, we also classified children into three categories: those who punished the in-group less than the out-group ( $N = 10$ ), those who punished the in-group more than the out-group ( $N = 23$ ), and those who punished the in-group and the out-group at the same degree ( $N = 39$ ). The three categories of children did not differ in their levels of in-group bias on the manipulation check (one-way ANOVA,  $p = 1.00$ ). Taken together, children's in-group biases on the manipulation check questions concerning in-group preference did not seem to be related to their responses toward different groups in the restoration and punishment task.



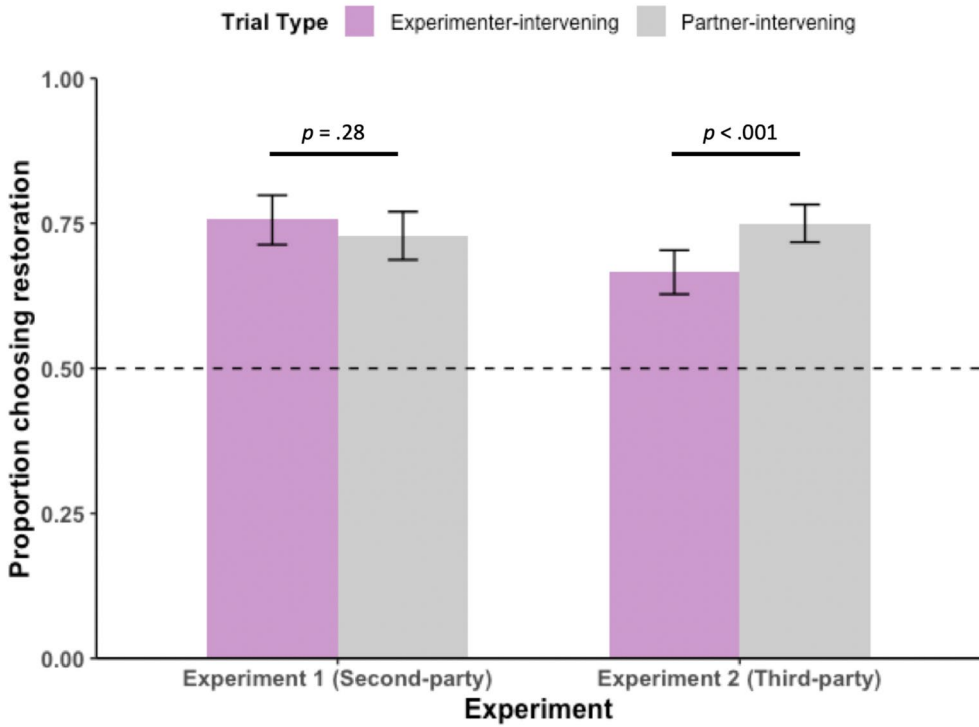
**FIGURE 4** Experiment 2: Forced-choice between punishment and restoration across trial types (experimenter-intervening and partner-intervening trials), groups (ingroup, outgroup, and unaffiliated partners), with two age groups (5 and 6 years) collapsed. Y-axis shows the proportion choosing restoration instead of punishment, ranging from 0 to 1. Error bars are bootstrapped 95% confidence intervals

### 3.2.3 | Comparing Experiment 1 (second-party) and Experiment 2 (third-party) results (age-matched samples)

To compare 5–6-year-olds' restoration versus punishment choices across second-party and third-party contexts, we further analyzed data in two studies (only including 5–6 years children in Experiment 1) predicting children's choices (0 = punishment, 1 = restoration) as a function of group (in-group, out-group, and unaffiliated), trial type (partner-intervening and experimenter-intervening), experiment (Experiment 1 and Experiment 2), and their interactions, with a random intercept for participant. The only significant effect was the interaction between trial type and experiment (see Figure 5),  $\beta = .83$ ,  $SE = .24$ , 95% CI [.36, 1.30], which was driven by the effect of trial type being significant in Experiment 2 but not in Experiment 1. Children favored restoration over punishment more in the partner-intervening trials than in the experimenter-intervening trials in Experiment 2 (third-party),  $\beta = .62$ ,  $SE = .15$ , 95% CI [.33, .92], but they did not choose these two options differently across trial types in Experiment 1 (second-party; but note that we found a significant group by trial type interaction there),  $\beta = -.20$ ,  $SE = .19$ , 95% CI [-.57, .17].

## 3.3 | Discussion

In Experiment 2, we invited Chinese children to a third-party game when their own welfare was not involved. Similar to the finding from second-party game in Experiment 1, we found that 5- to 6-year-olds chose restoration



**FIGURE 5** Experiment comparison: Interaction between trial type and Experiment (second-party Experiment 1 and third-party Experiment 2), averaged across groups and 5–6-year-old participants. Y-axis shows the proportion choosing restoration instead of punishment, ranging from 0 to 1. Error bars are bootstrapped 95% confidence intervals

more often than punishment. Importantly, children chose restoration more in the partner-intervening trials compared to the experimenter-intervening trials, suggesting strong preferences for restoration when witnessing a live transgression. In other words, children as young as 5 and 6 have a ‘victim-oriented’ focus when they are not personally involved. Compared to Experiment 1, as children’s own welfare was not affected here, these findings further imply a desire to return something to its rightful owner and a concern for protecting the victim as opposed to more harshly punishing the transgressor, consistent with the broader claim made by Riedl et al. (2015). Moreover, we were the first to show that this pattern persisted regardless of the partner’s group membership.

The relatively low levels of punishment as opposed to restoration again mirrored prior studies showing that children usually do not punish at above chance rates (e.g., Jordan et al., 2014; McAuliffe et al., 2015). These results also suggest that the punitive form of punishment might not be the preferred option in children. Future studies should consider alternative options like restoration when investigating punishment. Our findings are consistent with some adult work (Chavez & Bicchieri, 2013); we find that preschoolers favored restoration and that adult work adults compensated the victims more as compared to punishment in the third-party context (but see FeldmanHall et al., 2014 for somewhat different patterns).

#### 4 | GENERAL DISCUSSION

Across two experiments, we investigated Chinese preschoolers’ punishment and restoration behaviors in the face of transgressions both when their own welfare was involved (the second-party Experiment 1) and when it was not

(the third-party Experiment 2). We found that children by the age of 5 and 6 favored restoration over punishment, returning the object to its rightful owner without harshly punishing the transgressor across second- and third-party contexts. Interestingly, such a preference was more pronounced in the third-party context: children even increased restoration in partner-intervening trials (where they could punish the transgressor) compared to the experimenter-intervening trials (where they could not punish the experimenter who caused the unfair outcome, and thus, restoration was the expected choice, as per literature on children's sense of fairness, e.g., McAuliffe et al., 2017). We are the first to document such a preference for restoration in a collectivistic culture as well as a clear age effect: older children (5–6 years) favored restoration over punishment more than younger children (3–4 years), suggesting that children develop this preference during preschool years. Our finding is consistent with recent adult and developmental studies using different paradigms in Western individualistic cultures (e.g., Chavez & Bicchieri, 2013; FeldmanHall et al., 2014; Heffner & FeldmanHall, 2019; Riedl et al., 2015), and with literature on theories of law and restorative measures (e.g., Saulnier & Sivasubramaniam, 2015; Weitekamp, 1993). These findings dovetail with research in children's developing sense of fairness, as in this age range children increasingly favor fairness between individuals and generally show a preference for fair outcomes (Killen & Smetana, 2005). The age-related changes revealed in these studies also parallel theory of mind work (Wellman et al., 2006; Wimmer & Perner, 1983). We propose that the understanding of others' mental states might be a potential mechanism underlying the developmental changes here (for related work on how mental state understanding is related to fairness preferences and sharing behaviors: e.g., Takagishi et al., 2010; Tsoi & McAuliffe, 2019; Wu & Su, 2014). We discuss other potential motivations underlying children's restorative and punishment behaviors in the General Discussion.

Across two experiments, we assigned Chinese children to novel social groups using a color-preference/participant-choice assignment procedure and measured resulting in-group bias, following past work (Elashi & Mills, 2014; Gonzalez et al., 2020; Jordan et al., 2014; as well as a study that successfully used this same procedure in the same population, Wu & Gao, 2018). Although not our main focus, our study extends the minimal group effect in children to a non-WEIRD (Western, educated, individualistic, rich, and democratic) culture, where developmental work on this effect is relatively rare (though see Wu & Gao, 2018). We showed that children aged 3 to 6 favor their minimal in-group members and evaluated them more positively, replicating past work on children's minimal group bias in WEIRD societies (e.g., Dunham et al., 2011). We acknowledge that other assignment procedures (e.g., pure random assignment and assignment allegedly based on personality) might lead to different levels of bias, and future work could further explore these differences (but see Yang & Dunham, 2019 for a recent study that found the random assignment procedure led to equally strong in-group bias as more meaningful, personality-based groups).

Group membership significantly influenced choices in the second-party intervention context, but not in the third-party context. In the second-party context, compared to the experimenter-intervening trials, when puppets stole things children were more likely to punish out-group and unaffiliated puppets (i.e., they chose restoration less), but not in-group puppets. This result might imply in-group love but not out-group hate is driving the effect, which accords with and extends past research by adding a neutral, unaffiliated member (e.g., Baumgartner et al., 2012; Bernhard et al., 2006). By contrast, in the experimenter-intervening trials when the puppet did not intentionally harm the child participant's welfare but the experimenter created a disadvantage for the child and an advantage for the puppet, children were less likely to choose restoration (instead favoring punishment) with an in-group partner as compared to an out-group partner. Although we cannot definitively explain this effect, we suggest that the social comparison literature might provide some insight. In particular, children may have chosen the punishment option when they were motivated to receive more than their partners (for similar findings in children with different paradigms, see Blake & McAuliffe, 2011; Sheskin et al., 2014). In other words, they might compare themselves more with in-group members than out-group members because social comparisons happen more with closer and more similar others (Chafel, 1985; Festinger, 1954; Pettigrew, 2016; Smith et al., 2012;



Turner, 1975). This could motivate the desire to pursue a relative advantage in this case. Furthermore, people in collectivistic cultures, for example, China, have been shown to anticipate more threat and competition from in-groups (as compared to people in individualistic cultures like the United States; Liu et al., 2019), which might motivate their enhanced comparison motives with in-groups in our task. In addition to social comparison, other motivations such as fairness concerns (see McAuliffe et al., 2017 for a review), self- or joint-outcome maximization, and spite (McAuliffe et al., 2014) might also facilitate children's behaviors in the current study. Future studies could look at different motives behind punishment and restoration behavior and explore what role each of these motivations plays.

We note that the restoration option in our design (following Riedl et al., 2015) is different from the compensation option in other studies. In those studies (e.g., FeldmanHall et al., 2014; Heffner & FeldmanHall, 2019), the compensation option increases the victim's payoff by introducing extra resources whereas in our study, the restoration option restores the distribution to the original state ('returning the object to its rightful owner') and holds the total amount of resources constant. In both cases, however, the punishment option always reduces the perpetrator's payoff. Conceptually speaking, both designs are legitimate and commonplace: punishment options always focus on sanctioning the perpetrator whereas the compensation/restoration options always focus on protecting the victim. We believe both designs reveal interesting insights into norm enforcement and future studies will be needed to explore people's preferences among various restorative/compensatory options. That said, we note that our design has some advantages in terms of everyday realism, at least in the most common childhood contexts, because returning stolen property to a victim is likely more frequent than providing some additional compensation to victims while leaving the ill-gotten gains in the hands of a perpetrator. Future studies could compare these different designs as well as paradigms in the punishment literature (e.g., McAuliffe et al., 2017) to explore the shared and unique aspects of human cognition revealed by them.

It is also important to note that the punishment option in our study was designed to be a relatively severe form of punishment, in that the transgressor loses both of their own tokens in addition to the stolen one, thereby departing from Riedl and Colleagues (2015), in which the transgressor only loses the one object that they stole). As we discussed in the Introduction, we operationalized the punishment option this way to better convey the idea of punishment as deterrent—though the victim does not take the lost token back, the transgressor has to pay a bigger price for the wrongdoing, thus, potentially teaching them a lesson. However, this more severe form of punishment might have moved some children away from choosing punishment. Future studies could compare punishment options that differ in severity, for example, the transgressor losing one token instead of both tokens, and explore if children's relative preference between restoration and punishment changes as the punishment option become more or less severe.

Another contribution of the current work is the inclusion of a group-neutral/unaffiliated perpetrator in addition to in-group and out-group transgressors, something that past research usually lacks. The value of the neutral group is that it allows us to better pinpoint the direction of this bias if it exists—i.e., is it driven by in-group love (i.e., preferential treatment of in-groups), out-group hate (i.e., discriminatory treatment of out-groups), or both? As far as we know, there is only one adult study that includes an unaffiliated violator (Schiller et al., 2014), and it finds that bias in third-party punishment is driven by both in-group love and out-group hate. In the current research, we explored this issue in children by including a group-unaffiliated individual (following Schiller et al., 2014, we introduced it as someone who had not decided about which group to join and thus, it had not joined either group). It thus allows us to further understand how intergroup bias affects norm enforcement, and whether it mirrors the developmental trajectory of intergroup attitudes (Buttelmann & Böhm, 2014; Nesdale, 2004). Unexpectedly, results for the neutral target did not always fall in between the in-group and the out-group. It is possible that children saw this undecided puppet as a complete outlier or outsider, and therefore, treated them even more harshly (i.e., perhaps via wondering why everyone belonged to group except this puppet). More research using neutral targets with children (as well as with different ways of introducing the neutral target) is needed to explore how children

reason about the neutral/unaffiliated members and to more comprehensively investigate the directionality of intergroup bias.

We also acknowledge that across manipulation check questions and main behavioral trials we always pitted two options, for example, in-group versus out-group or restoration versus punishment, against each other without providing other options like 'both' or 'neither'. This use of dichotomous, forced-choice measures is common in the developmental literature, but we acknowledge that it might restrict children's responses and it will be important to examine how children might respond if other options are available. Additionally, our main behavioral trials involved decisions around unnecessary resources (stickers) as compared to necessary ones (such as medical or school supplies, which might create stronger concerns for equal distributions, and thus, reduce the possibility of finding in-group bias, see Elenbaas & Killen, 2016; Rizzo et al., 2016). A fruitful avenue of future work might be the investigation of children's restoration and punishment behavior across different types of resources.

In summary, we showed that preschoolers preferred restoration to punishment, and that they emphasized restorative justice not only for themselves, but also for others (i.e., an in-group member) when their self-interest was not involved. There was an age-related increase in favoring restoration between the age of 3 and 6 (in second-party games). Children's responses were also shaped by intergroup concerns in complex ways. Our research challenges the view that punishment is the standard response to transgressions, and suggests that alternative options like restoration might instead be preferred by even young children. Future research is needed to explore the preference for restoration and how it is influenced by intergroup bias across development and cultures.

## ACKNOWLEDGEMENTS

We are grateful to Yang Gao and Liyueer Baxueyuan Kindergarten, Honghuanglan Kindergarten, Aisichule Preschool, and Beidaxueyuan Kindergarten for help with participant recruitment; and to all children for their participation. We thank Xin Liu, Xiuyan Li, and Fan Wang for assistance in data collection; Qiuchen Ma for being a second coder; and members of the Child Development Lab at Tsinghua University for helpful discussions. This research was supported by grants from the National Natural Science Foundation of China (31600899), Tsinghua University Initiative Scientific Research Program (20171080349), Tsinghua University Tang Scholar Program, and Tsinghua University Lifelong Learning Lab to ZW, as well as Tsinghua University Undergraduate Research Grant to XY.

## ETHICS STATEMENT

The ethics committee of Tsinghua University approved the study, and the treatment of human participants was in accord with the ethical standards of the American Psychological Association; the rights of the participants were protected and applicable human research participant guidelines were followed.

## CONFLICT OF INTEREST

The authors declare that they have no competing interests.

## DATA AVAILABILITY STATEMENT

Data and analysis code that replicate all results in the manuscript and Supporting Information are also openly shared and are available on an anonymous link at: [https://osf.io/87hme/?view\\_only=782b53ad77e94ea1a2b6dd2caee936c8](https://osf.io/87hme/?view_only=782b53ad77e94ea1a2b6dd2caee936c8).

## ORCID

Xin Yang  <https://orcid.org/0000-0002-7506-9673>

Zhen Wu  <https://orcid.org/0000-0003-3655-1069>

Yarrow Dunham  <https://orcid.org/0000-0002-4265-4438>

## REFERENCES

- Abrams, D., & Killen, M. (2014). Social exclusion of children: Developmental origins of prejudice. *Journal of Social Issues*, 70(1), 1–11. <https://doi.org/10.1111/josi.12043>
- Abrams, D., Rutland, A., Pelletier, J., & Ferrell, J. M. (2008). Children's judgments of disloyal and immoral peer behavior: Subjective group dynamics in minimal intergroup contexts. *Child Development*, 79(2), 444–461. <https://doi.org/10.1111/j.1467-8624.2007.01135.x>
- Barret, M. (2007). *Children's knowledge, beliefs and feelings about nations and national groups*. Psychology Press.
- Baumgartner, T., Götte, L., Gügler, R., & Fehr, E. (2012). The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. *Human Brain Mapping*, 33(6), 1452–1469. <https://doi.org/10.1002/hbm.21298>
- Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, 442, 912–915. <https://doi.org/10.1038/nature04981>
- Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child Development*, 68(3), 530–543. <https://doi.org/10.1111/j.1467-8624.1997.tb01956.x>
- Blake, P. R., & McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition*, 120(2), 215–224. <https://doi.org/10.1016/j.cognition.2011.04.006>
- Buttelmann, D., & Böhm, R. (2014). The ontogeny of the motivation that underlies in-group bias. *Psychological Science*, 25(4), 921–927. <https://doi.org/10.1177/0956797613516802>
- Chafel, J. A. (1985). Social comparison and the young child: Current research issues. *Early Child Development and Care*, 21(1–3), 35–59. <https://doi.org/10.1080/0300443850210104>
- Chavez, A. K., & Bicchieri, C. (2013). Third-party sanctioning and compensation behavior: Findings from the ultimatum game. *Journal of Economic Psychology*, 39, 268–277. <https://doi.org/10.1016/j.joep.2013.09.004>
- Dunham, Y. (2018). Mere membership. *Trends in Cognitive Sciences*, 22(9), 780–793. <https://doi.org/10.1016/j.tics.2018.06.004>
- Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development*, 82(3), 793–811. <https://doi.org/10.1111/j.1467-8624.2011.01577.x>
- Elashi, F. B., & Mills, C. M. (2014). Do children trust based on group membership or prior accuracy? The role of novel group membership in children's trust decisions. *Journal of Experimental Child Psychology*, 128, 88–104. <https://doi.org/10.1016/j.jecp.2014.07.003>
- Elenbaas, L., & Killen, M. (2016). Children rectify inequalities for disadvantaged groups. *Developmental Psychology*, 52(8), 1318–1329. <https://doi.org/10.1037/dev0000154>
- Falk, C. F., Heine, S. J., & Takemura, K. (2014). Cultural variation in the minimal group effect. *Journal of Cross-Cultural Psychology*, 45(2), 265–281. <https://doi.org/10.1177/0022022113492892>
- Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in Cognitive Sciences*, 8(4), 185–190. <https://doi.org/10.1016/j.tics.2004.02.007>
- FeldmanHall, O., Sokol-Hessner, P., Van Bavel, J. J., & Phelps, E. A. (2014). Fairness violations elicit greater punishment on behalf of another than for oneself. *Nature Communications*, 5, 5306. <https://doi.org/10.1038/ncomms6306>
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140. <https://doi.org/10.1177/001872675400700202>
- Gonzalez, G., Blake, P. R., Dunham, Y., & McAuliffe, K. (2020). Ingroup bias does not influence inequity aversion in children. *Developmental Psychology*, 56(6), 1080–1091. <https://doi.org/10.1037/dev0000924>
- Heffner, J., & FeldmanHall, O. (2019). Why we don't always punish: Preferences for non-punitive responses to moral violations. *Scientific Reports*, 9(1), 1–13. <https://doi.org/10.1038/s41598-019-49680-2>
- Jordan, J. J., McAuliffe, K., & Warneken, F. (2014). Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences*, 111(35), 12710–12715. <https://doi.org/10.1073/pnas.1402280111>
- Killén, M., & Smetana, J. (Eds.) (2005). *Handbook of moral development*. Psychology Press.
- Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, 104(30), 12577–12580. <https://doi.org/10.1073/pnas.0705345104>
- Kubota, J. T., Li, J., Bar-David, E., Banaji, M. R., & Phelps, E. A. (2013). The price of racial bias: Intergroup negotiations in the Ultimatum Game. *Psychological Science*, 24(12), 2498–2504. <https://doi.org/10.1177/0956797613496435>
- Liu, S. S., Morris, M. W., Talhelm, T., & Yang, Q. (2019). Ingroup vigilance in collectivistic cultures. *Proceedings of the National Academy of Sciences*, 116(29), 14538–14546. <https://doi.org/10.1073/pnas.1817588116>
- McAuliffe, K., Blake, P. R., Steinbeis, N., & Warneken, F. (2017). The developmental foundations of human fairness. *Nature Human Behaviour*, 1, 0042. <https://doi.org/10.1038/s41562-016-0042>
- McAuliffe, K., Blake, P. R., & Warneken, F. (2014). Children reject inequity out of spite. *Biology Letters*, 10(12), 20140743. <https://doi.org/10.1098/rsbl.2014.0743>

- McAuliffe, K., & Dunham, Y. (2016). Group bias in cooperative norm enforcement. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1686), 20150073. <https://doi.org/10.1098/rstb.2015.0073>
- McAuliffe, K., & Dunham, Y. (2017). Fairness overrides group bias in children's second-party punishment. *Journal of Experimental Psychology: General*, 146(4), 485.
- McAuliffe, K., Jordan, J. J., & Warneken, F. (2015). Costly third-party punishment in young children. *Cognition*, 134, 1–10. <https://doi.org/10.1016/j.cognition.2014.08.013>
- McLeish, K. N., & Oxoby, R. J. (2011). Social interactions and the salience of social identity. *Journal of Economic Psychology*, 32(1), 172–178. <https://doi.org/10.1016/j.joep.2010.11.003>
- Mendoza, S. A., Lane, S. P., & Amodio, D. M. (2014). For members only: Ingroup punishment of fairness norm violations in the ultimatum game. *Social Psychological and Personality Science*, 5(6), 662–670. <https://doi.org/10.1177/1948550614527115>
- Nesdale, D. (2004). Social identity and ethnic prejudice. In M. Bennett & F. Sani (Eds.), *The development of the social self* (pp. 219–245). Psychology Press.
- Pettigrew, T. F. (2016). In pursuit of three theories: Authoritarianism, relative deprivation, and intergroup contact. *Annual Review of Psychology*, 67(1), 1–21. <https://doi.org/10.1146/annurev-psych-122414-033327>
- Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science*, 24(6), 999–1006. <https://doi.org/10.1177/0956797612466267>
- Richter, N., Over, H., & Dunham, Y. (2016). The effects of minimal group membership on young preschoolers' social preferences, estimates of similarity, and behavioral attribution. *Collabra Psychology*, 2(1), 8. <https://doi.org/10.1525/collabra.44>
- Riedl, K., Jensen, K., Call, J., & Tomasello, M. (2015). Restorative justice in children. *Current Biology*, 25(13), 1731–1735. <https://doi.org/10.1016/j.cub.2015.05.014>
- Rizzo, M. T., Elenbaas, L., Cooley, S., & Killen, M. (2016). Children's recognition of fairness and others' welfare in a resource allocation task: Age related changes. *Developmental Psychology*, 52(8), 1307–1317. <https://doi.org/10.1037/dev0000134>
- Saulnier, A., & Sivasubramaniam, D. (2015). Restorative justice: Underlying mechanisms and future directions. *New Criminal Law Review: An International and Interdisciplinary Journal*, 18(4), 510–536. <https://doi.org/10.1525/nclr.2015.150014>
- Schiller, B., Baumgartner, T., & Knoch, D. (2014). Intergroup bias in third-party punishment stems from both ingroup favoritism and outgroup discrimination. *Evolution and Human Behavior*, 35(3), 169–175. <https://doi.org/10.1016/j.evolhumbehav.2013.12.006>
- Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, 124(3), 325–333. <https://doi.org/10.1016/j.cognition.2012.06.004>
- Sheskin, M., Bloom, P., & Wynn, K. (2014). Anti-equality: Social comparison in young children. *Cognition*, 130(2), 152–156. <https://doi.org/10.1016/j.cognition.2013.10.008>
- Smith, H. J., Pettigrew, T. F., Pippin, G. M., & Bialosiewicz, S. (2012). Relative deprivation: A theoretical and meta-analytic review. *Personality and Social Psychology Review*, 16(3), 203–232. <https://doi.org/10.1177/1088868311430825>
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223(5), 96–102. <https://doi.org/10.1038/scientificamerican1170-96>
- Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., & Yamagishi, T. (2010). Theory of mind enhances preference for fairness. *Journal of Experimental Child Psychology*, 105(1–2), 130–137. <https://doi.org/10.1016/j.jecp.2009.09.005>
- Tsoi, L., & McAuliffe, K. (2019). Individual differences in theory of mind predict inequity aversion in children. *Personality and Social Psychology Bulletin*, 46(4), 559–571. <https://doi.org/10.1177/0146167219867957>
- Turner, J. C. (1975). Social comparison and social identity: Some prospects for intergroup behaviour. *European Journal of Social Psychology*, 5(1), 1–34. <https://doi.org/10.1002/ejsp.2420050102>
- Valenzuela, A., & Srivastava, J. (2012). Role of information asymmetry and situational salience in reducing intergroup bias: The case of ultimatum games. *Personality and Social Psychology Bulletin*, 38(12), 1671–1683. <https://doi.org/10.1177/0146167212458327>
- Weitekamp, E. (1993). Reparative justice: Towards a victim oriented system. *European Journal on Criminal Policy and Research*, 1(1), 70–93. <https://doi.org/10.1007/BF02249525>
- Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological Science*, 17(12), 1075–1081. <https://doi.org/10.1111/j.1467-9280.2006.01830.x>
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)
- Wu, Z., & Gao, X. (2018). Preschoolers' group bias in punishing selfishness in the Ultimatum Game. *Journal of Experimental Child Psychology*, 166, 280–292. <https://doi.org/10.1016/j.jecp.2017.08.015>
- Wu, Z., & Su, Y. (2014). How do preschoolers' sharing behaviors relate to their theory of mind understanding? *Journal of Experimental Child Psychology*, 120, 73–86. <https://doi.org/10.1016/j.jecp.2013.11.007>

- Yang, X., & Dunham, Y. (2019). Minimal but meaningful: Probing the limits of randomly assigned social identities. *Journal of Experimental Child Psychology*, 185, 19–34. <https://doi.org/10.1016/j.jecp.2019.04.013>
- Yee, M., & Brown, R. (1994). The development of gender differentiation in young children. *British Journal of Social Psychology*, 33(2), 183–196. <https://doi.org/10.1111/j.2044-8309.1994.tb01017.x>
- Yudkin, D. A., Van Bavel, J. J., & Rhodes, M. (2020). Young children police in-group members at personal cost. *Journal of Experimental Psychology: General*, 149(1), 182. <https://doi.org/10.31234/osf.io/rc4ta>

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

Supplementary Material

**How to cite this article:** Yang X, Wu Z, Dunham Y. Children's restorative justice in an intergroup context. *Social Development*. 2021;00:1–21. doi: 10.1111/sode.12508.