Research Article

Psychological Science
2024, Vol. 35(10) 1094–1107

DOI: 10.1177/09567976241267854
www.psychologicalscience.org/PS

Children Sustain Cooperation in a Threshold Public-Goods Game Even When Seeing Others' Outcomes

Patricia Kanngiesser[1,2], Jahnavi Sunderarajan[2], Sebastian Hafenbrädl[3], and Jan K. Woike[1,4]

[1]School of Psychology, University of Plymouth; [2]Faculty of Education and Psychology, Freie Universität Berlin; [3]Department of Managing People in Organizations, IESE Business School; and [4]Center for Adaptive Rationality, Max Planck Institute for Human Development

Abstract

Many societal challenges are threshold dilemmas requiring people to cooperate to reach a threshold before group benefits can be reaped. Yet receiving feedback about others' outcomes relative to one's own (*relative feedback*) can undermine cooperation by focusing group members' attention on outperforming each other. We investigated the impact of relative feedback compared to *individual feedback* (only seeing one's own outcome) on cooperation in children from Germany and India (6- to 10-year-olds, $N = 240$). Using a threshold public-goods game with real water as a resource, we show that, although feedback had an effect, most groups sustained cooperation at high levels in both feedback conditions until the end of the game. Analyses of children's communication (14,374 codable utterances) revealed more references to social comparisons and more verbal efforts to coordinate in the relative-feedback condition. Thresholds can mitigate the most adverse effects of social comparisons by focusing attention on a common goal.

Keywords

cooperation, social dilemma, social comparisons, development, cross-cultural, open data, open materials

Social dilemmas underlie many pressing societal challenges. Although societies would benefit if everyone cooperated to address these challenges, individual actors are better off by free riding and letting others pay the cost of cooperation (Dawes, 1980; Kollock, 1998; van Lange et al., 2013). Cooperative behavior in social dilemmas increases when groups need to reach a threshold before they can reap benefits (Cadsby & Maynes, 1999; Deutchman et al., 2022; van de Kragt et al., 1983). In fact, many real-world problems resemble threshold dilemmas (Barrett, 2007; Milinski et al., 2008). For example, smallpox was eradicated in 1980 after decades of work resulting in sufficient resources for vaccination and elimination in every country (Barrett, 2007). Moreover, thresholds also feature prominently in calls to mitigate the most disastrous consequences of human-made climate change (Lee et al., 2023). Thresholds can act as strong focal points that transform social dilemmas into coordination problems (Barrett & Dannenberg, 2012, 2014; Cadsby & Maynes, 1999; Dannenberg et al., 2015).

Cooperative behavior, however, can be undermined when group members focus on outperforming each other (Garcia et al., 2013). Such a shift in focus can result from receiving feedback about everyone's personal outcome (relative feedback), which allows for relative comparisons and can draw attention to disparities in gains (McClintock & Nuttin, 1969; Messick & McClintock, 1968). Relative feedback can encourage free riding to maximize differences in outcomes and reduce cooperation rates in nonthreshold social-dilemma games (McClintock & McNeel, 1966; Messick & McClintock, 1968; Woike & Hafenbrädl, 2020). It could have similar detrimental effects in threshold

Corresponding Author:
Patricia Kanngiesser, School of Psychology, University of Plymouth
Email: patricia.kanngiesser@plymouth.ac.uk

games, but thresholds could also mitigate these effects by focusing attention on a common goal. To our knowledge, no study has investigated this to date.

It is particularly interesting to investigate how outcome feedback impacts children's cooperation in threshold games. Young children are already able to tacitly and explicitly coordinate their actions with others and enter joint obligations to achieve mutually beneficial outcomes (Tomasello, 2022). In social dilemmas, children in diverse societies begin to cooperate from age 5 to 6 years (Alencar et al., 2008; Angerer et al., 2016; Dutra et al., 2018; Harbaugh & Krause, 2000; Houser et al., 2012; Prétôt & McAuliffe, 2020; Vogelsang et al., 2014). For example, German 6-year-olds cooperated to maintain a common pool resource (water) that would collapse if water levels fell below a threshold (Koomen & Herrmann, 2018). Similarly, German 6- to 11-year-olds sustained a fishing resource for longer when withdrawal was limited (Ebersbach et al., 2019). When failing to reach a threshold would have resulted in a certain loss of endowments, up to 90% of groups of 6- to 10-year-olds from the United States, China, and the Democratic Republic of Congo cooperated (Bowie et al., 2022). This suggests that children are able to cooperate and that thresholds facilitate children's cooperation.

From about 5 years of age, children compare themselves to others (Butler, 1989; Steinbeis & Singer, 2013). Importantly, children in diverse societies are sensitive to feedback about others' outcomes in middle childhood (McClintock, 1974; McClintock et al., 1977; McClintock & Moskowitz, 1976). For example, children made more competitive choices in a maximizing-differences game when seeing their own and others' outcomes compared to seeing only their own outcomes (McClintock & Nuttin, 1969; Toda et al., 1978). It is an open question whether outcome feedback will similarly affect children's cooperation in a threshold public-goods game. Middle childhood seems the ideal developmental period to study this, because both cooperative abilities and comparisons with others emerge at that age.

In our study, groups of 6- to 10-year-olds played a novel threshold public-goods game over eight rounds. Our game used a continuous-contribution mechanism that prevented others from observing the magnitude of individual contributions during a round and thus provided opportunities for free riding without being noticed. To make the game tangible and engaging, we used real water as a resource (see Fig. 1). Each child received a water bottle at the start of each round (containing 400 ml water), which was connected to a pool with an artificial animal. The animal was described as thirsty and could only drink if the water reached a threshold (600 ml). Each child (privately) decided how much water to contribute by pressing a tap on the bottle. At least two children needed to contribute to fill the pool to the threshold (see Fig. 2). If the threshold

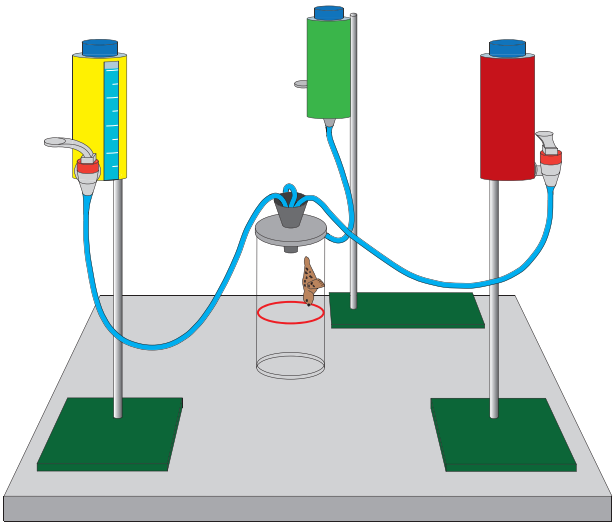

Fig. 1. Schematic of the water game. Each child sat in front of a water bottle filled with 400 ml water at the start of each round. Each bottle had a tap that children could use to release water into the central pool (transparent container). Bottles were covered in the child's assigned color, and the water level in the bottle was only visible to the child sitting in front of it. The threshold in the pool (600 ml) was marked with a red line, and the "thirsty" animal (a seal) was placed just above it. All water tubes led to a funnel above the pool, obscuring which child was contributing water.

Statement of Relevance

Many real-world problems require people to cooperate, but individuals are often better off by free riding and letting others pay the cost of cooperation. Feedback about own and others' outcomes can undermine cooperation by focusing attention on disparities in gains and encouraging people to outperform each other. However, thresholds that a group needs to reach before it can reap group benefits could attenuate this effect by focusing attention on a common goal. We developed a novel game for children with real water as resource that children could donate to reach a threshold. Six- to 10-year-olds from Germany and India sustained cooperation at high levels, even when provided with feedback about their own and others' outcomes. Analyses of children's communication revealed broad similarities but also nuanced societal variation in line with prevalent cultural models. Clear and unambiguous thresholds can foster cooperation and provide a buffer against the most detrimental effects of social comparisons.

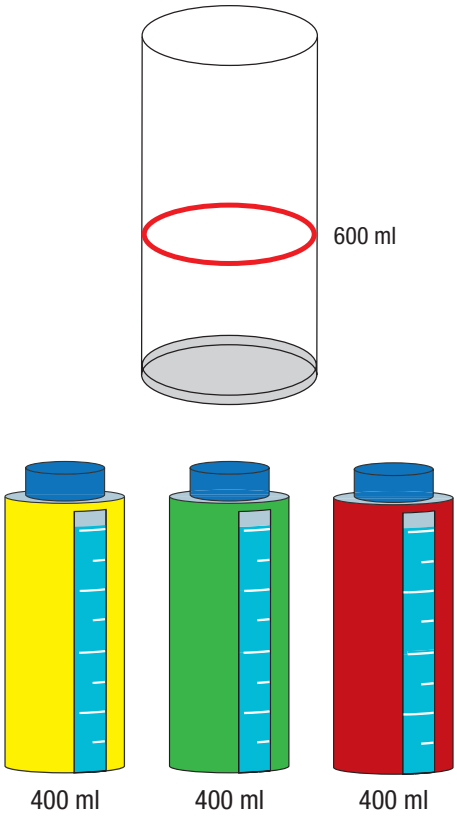

Fig. 2. Schematic of the threshold game. To reach the threshold, children needed to add 600 ml of water to the pool. Because each bottle contained 400 ml of water, at least two children needed to donate water to fill the pool to the threshold.

was reached, the animal rewarded each child with a water bag (350 ml water) that they could keep in addition to any remaining water in their respective bottles. To incentivize the game, we awarded prizes proportionate to the amount of water children had collected.

The remaining water in children's bottles was collected in large tubes (see Fig. 3). After each round, children could either visually compare how much water everyone in the group had collected (relative-feedback condition; Fig. 3a) or observe how much water they had individually collected (individual-feedback condition, between-subjects condition; Fig. 3b). In the relative-feedback condition, differences in outcomes accumulated and become increasingly noticeable over time. If feedback about everyone's outcomes focuses children's attention on their relative position within the group and increases free riding (McClintock & McNeel, 1966; McClintock & Nuttin, 1969; Toda et al., 1978; Woike & Hafenbrädl, 2020), then we would expect a stronger decrease in cooperation over successive rounds in the relative-feedback condition compared to the individual-feedback condition (Condition × Round interaction).

In middle childhood, cultural variation in cooperative behaviors (Blake et al., 2015; Bowie et al., 2022; House et al., 2013, 2020) but also cheating (Kanngiesser et al., 2023) and competitive behaviors (McClintock, 1974; McClintock & Nuttin, 1969; Toda et al., 1978) becomes pronounced. To study potential cultural variation in children's susceptibility to feedback in our threshold game, we conducted our study in Germany and India. The Indian education system is highly competitive, with grades significantly determining future opportunities and high achievers often celebrated (Boucher, 2024). Indian children may be more susceptible to feedback about their relative position in a group than German children, and if this were the case, we would expect to find an interaction between condition, round, and location.

Finally, children could freely talk to each other during the game, allowing us to systematically investigate their communicative strategies and their understanding of the game. Studies with adults have shown that cooperation increases when people can communicate (Balliet, 2010; Sally, 1995), but not all communication in children is equally effective in fostering cooperation (Grueneisen & Tomasello, 2017, 2019; Koomen & Herrmann, 2018). Our coding scheme included verbal coordination strategies (e.g., giving water, goals), use of normative language (Grueneisen & Tomasello, 2017, 2019), promises and threats (Ostrom et al., 1992; Sally, 1995), as well as references to fairness, reciprocity, and other relevant utterances (see the Supplemental Material available online for the full coding scheme). Importantly, we also coded whether children uttered social comparisons (e.g., talking about how much water they and others had). If relative feedback about everyone's outcomes focuses attention on disparities in outcomes, we would expect children to utter more social comparisons in the relative-feedback condition than in the individual-feedback condition. Relatedly, if relative feedback made it more difficult to sustain cooperation, we would expect children to increase their verbal efforts to coordinate water contributions. Our study thus provides insight into German and Indian children's behavioral and communicative strategies when solving a novel water-threshold public-goods game and the effect of different types of outcome feedback on children's cooperation.

Open Practices Statement

No aspects of the study were preregistered. The study materials, the data underlying the analyses reported in the article, and the analysis script are publicly available (https://osf.io/xzfce/). The study materials are also available in the Supplemental Material. Transcripts of

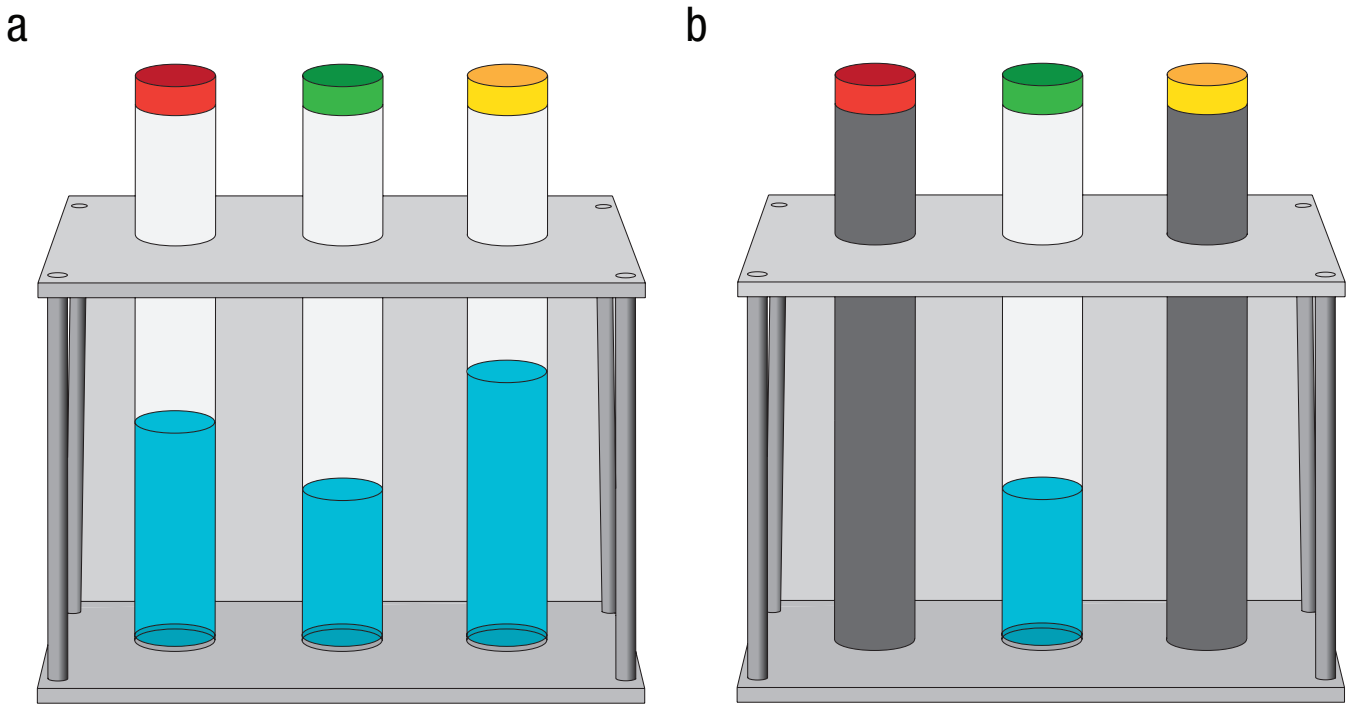

Fig. 3. Schematic of the water collection tubes in the two conditions (between subjects). Each child emptied any remaining water into a collection tube, marked with the child's assigned color. In the relative-feedback condition, children observed the water levels in all collection tubes and could thus compare their own outcomes to those of others in the group. In the individual-feedback condition, children could only observe their own water level, and the water levels in the other tubes were occluded with removable covers. The relative-feedback condition is illustrated in (a) and the individual-feedback condition in (b).

children's conversations are available on request to the corresponding author.

Method

Participants

The final sample consisted of 240 children, aged 6 to 10 years: 117 children (39 groups) from Berlin, Germany (M = 7.2 years, SD = 1.1 years; 57 female, or 49%) and 123 children (41 groups) from Pune, India (M = 8.1 years, SD = 1.3 years; 63 female, or 51%). Children participated in the study in groups of three, matched by gender and age (a 1-year age gap was permissible), resulting in a total of 80 groups. Half of the groups were assigned to the individual-feedback condition (Germany: 19 groups; India: 21 groups) and half to the relative-feedback condition (Germany: 20 groups; India: 20 groups; between-subjects design), with gender and age balanced across conditions (for further details, see the Supplemental Material). The sample size was based on previous studies (Koomen & Herrmann, 2018; Woike & Hafenbrädl, 2020).

We excluded one group (Germany) because one child was picked up from the after-school program during the study, and the group completed only four rounds in total. Two additional German groups had missing data for one of the rounds (Round 8 and Round 7, respectively) because of interruptions. Data from these two groups were included in the main analyses. One of the groups (with Round 7 data missing) was accidentally not transcribed; hence, no verbal coding is available for this group. For information on participants who took part in the pilot study, see the Supplemental Material.

Ethics and consent

The study was approved by the Ethics Committee of the Faculty of Education and Psychology at Freie Universität Berlin (Approval No. 170/2018). In order to recruit and test in local schools, we also received approval from the Berlin School Senate. Only children whose parents had consented to their child's participation took part in the study. Children were asked for their assent prior to participating in the game.

Additional details about method and materials, including the script used, are available in the Supplemental Material.

Materials

Children sat at a table about equidistant from each other. Each child was assigned a color at the start of the game (red, green, yellow) and sat in front of a laboratory stand with a water bottle, wrapped in the child's color, at approximately eye level. The water level in a bottle was only visible to the child sitting right in front of it (see Fig. 1). Each bottle contained 400 ml water at the start of each round. Each bottle had a plastic water tap and was connected to a central pool in the middle of the table through a plastic tube. The central pool included a marker for the threshold (600 ml) and a plastic toy animal (a seal) near the threshold. We used a large, sand-filled hourglass to indicate the remaining time in each round (90 s).

Children collected any leftover water from their bottles in large Plexiglas tubes; these were placed on a separate table and marked in children's assigned colors. During the rounds, the water levels in the tubes were not visible and only became observable (from behind the table) by turning the cover around the tube. This allowed us to manipulate (between subjects) whether children observed only their own accumulated water (the individual-feedback condition) or both their own and others' accumulated water (the relative-feedback condition). When children reached the threshold in the central pool, they each received a 350-ml water bag, which they collected in plastic containers placed in front of the water-collection tubes (because each child received a water bag, water bags did not contribute to relative differences in outcome between children). We displayed round numbers with a magnetic board showing the number of the current round and the number of remaining rounds.

Procedure

Two experimenters (female) conducted the study in a quiet room in children's schools, with one experimenter (Experimenter 1) instructing children and the other experimenter (Experimenter 2) assisting with running the study. The procedure was divided into two phases: a warm-up phase and an experimental phase.

Warm-up. Experimenter 1 first demonstrated the water-game apparatus, and children learned how to release water from their bottle and how to observe their bottle's water level. Experimenter 1 also introduced the central pool, pointing out the threshold and the thirsty animal. To help children understand how much water would be needed to reach the threshold, Experimenter 1 demonstrated what would happen if one person tried to fill the pool alone, two people tried to fill the pool, or three people tried to fill the pool. Experimenter 1 also introduced the hourglass that indicated how much time was left in a round.

Main phase. Next, Experimenter 1 told children that they would play the game over eight rounds and would decide in each round how much water to release into the central pool. Experimenter 1 also explained that each child would keep the leftover water in his or her bottle and collect it in their collection tube. If the group reached the threshold, the animal would be happy and reward each child with an additional water bag of 350 ml. Experimenter 1 further explained that each child would receive rewards proportionate to how much water had been collected (i.e., participants who collected a lot of water would receive many prizes), using a schematic to visualize the relation between water and prizes. Experimenter 1 also demonstrated the water-collection tubes. Depending on condition, children either saw only the collected water in their own tube (individual-feedback condition) or observed the water in everyone's tubes (relative-feedback condition).

Each round started with Experimenter 1 announcing the round number, turning around the hourglass, and leaving the room with Experimenter 2. Once the time was up, they reentered the room. If children had successfully reached the threshold, each child received a water bag. Children then poured the remaining water from their bottles into measuring cups for weighing (piloting had revealed that weighing water was the most accurate and efficient way to determine how much water remained in each child's bottle). Once each child's water was weighed, the water was poured into each child's collection tube.

In the individual-feedback condition, each child was privately shown how much water they had collected in their collection tube, ensuring this was not visible to the other children. In the relative-feedback condition, all three children were shown how much water everyone had collected in their respective tubes.

At the end of the game, Experimenter 1 and Experimenter 2 measured the water in the collection tubes. Children were told that they would receive their prizes once all children in their school had played the game. We used stationery items (pencils, erasers, etc.) as prizes. We deliberately converted collected water into prizes and did not tell children what the prizes would be to minimize potential issues around differences in the perceived value of the actual prizes. Reward equivalency has been discussed as a challenge, particularly for cross-cultural studies, and in recent years researchers have opted to use token or token-like resources (later converted into prizes) for their studies to circumvent this issue (Bowie et al., 2022; House et al., 2020).

For a detailed description of the procedure and the study script, see the Supplemental Material.

Data recording, coding, and analyses

After each round, we measured the remaining water in each bottle and the amount of water in the central pool using a kitchen scale (1 ml = 1 g). At the end of the game, we also measured the amount of water in each collection tube. Two German groups had missing data for one round (Rounds 7 and 8, respectively) because of interruptions. These groups were included in data analyses, with missing rounds coded as N/A. We checked the accuracy of our measurement by comparing the sum of individual measures per round to the final amount of water in the collection tube. There was a deviation in Germany of $M = 1.07\%$, $SD = 2.68\%$, and a deviation in India of $M = 1.49\%$, $SD = 2.77\%$, indicating very good accuracy.

We recorded children with two cameras, and transcribed (and, if necessary, translated) their conversations during each round of the game. We accidentally did not transcribe utterances of one German group (the group with missing data for Round 7 water contributions), and hence verbal coding is missing for this group. We developed a detailed coding scheme for children's utterances, and one coder scored utterances of German and Indian children. Across locations we transcribed 16,165 utterances. Of these 14,374 (89%) were scored as codable and 1,791 (11%) as noncodable (i.e., were inaudible or incomprehensible). There were somewhat more codable utterances in Germany (8,203, or 94%) than in India (6,171, or 83%). A second coder scored 25% of groups (3,587 utterances)—balanced by location, age, and condition—for reliability purposes. Reliability was good to excellent for the most frequently coded categories (i.e., >50 scored utterances), $\kappa \geq .71$.

All data analyses were conducted in R (R Core Team, 2023). We used generalized linear mixed models (GLMMs), as implemented in the *lme4* package (Bates et al., 2015), and conducted model comparisons to determine the model with the best fit to the data. To illustrate model specifications, we compared the following nested models to analyze whether groups reached the threshold:

$$\text{full: threshold} \sim \text{location} \times \text{condition} \times \text{round} + \text{gender} + \text{age} + (1 \mid \text{group\_ID}) \quad (1)$$

$$\text{red: threshold} \sim \text{location} \times \text{condition} + \text{location} \times \text{round} + \text{condition} \times \text{round} + \text{gender} + \text{age} + (1 \mid \text{group\_ID}) \quad (2)$$

$$\text{main: threshold} \sim \text{location} + \text{condition} + \text{round} + \text{gender} + \text{age} + (1 \mid \text{group\_ID}) \quad (3)$$

$$\text{null: threshold} \sim (1 \mid \text{group\_ID}) \quad (4)$$

For fixed effects in the best-fit model, *p* values were determined using likelihood ratio tests. To plot model estimates, we used the *plot_model* function in the *sjPlot* package (Lüdecke, 2023). Kappa values for reliability coding were calculated using the *irr* package (Gamer et al., 2019). Deidentified data and the R code to reproduce analyses and figures can be found on the Open Science Framework: https://osf.io/xzfce/.

Results

Threshold

Reaching the threshold. The majority of groups in both countries and conditions reached the threshold in all rounds of the game (see Fig. 4 and the Supplemental Material), even in the last round of the relative-feedback condition (India: 65%; Germany: 75%). To analyze groups' behavior, we fitted generalized linear mixed models (GLMMs) and included reaching of threshold (yes/no) as a response variable and feedback condition (individual vs. relative), location (Germany vs. India), round (1 to 8), gender, and average age of group members as fixed effects. Group number was entered as a random intercept (see the Method section for further details). We conducted model comparisons starting with a three-way interaction model of Location × Condition × Round to test for variation in threshold attainment over rounds in the two locations (age and gender were included as main effects).

The full three-way interaction model had a significantly better fit to the data than a null model containing only the random intercept, $\chi^2(9) = 35.94$, $p < .001$. Further model comparisons revealed that the model with the best fit to the data was the two-way interaction model (full vs. two-way: $\chi^2(1) = 2.28$, $p = .13$; two-way vs. main-effects model: $\chi^2(3) = 12.81$, $p = .005$). Likelihood ratio tests showed a significant Condition × Round interaction, *Estimate* (*Est*) = 0.43, $SE = 0.14$, $\chi^2(1) = 10.71$, $p = .001$: The number of groups that reached the threshold decreased more in the relative-feedback condition than in the individual-feedback condition (see Fig. 4; for a plot of the raw data, see the Supplemental Material). There was also a significant effect of gender, $Est = 1.05$, $SE = 0.51$, $\chi^2(1) = 4.22$, $p = .040$, with female groups being somewhat more successful in reaching the threshold than male groups. No other factors were significant (see Table 1 for details).

Conditional cooperation. In exploratory analyses, we also tested whether children cooperated conditionally by adding the results from the previous round (threshold reached = yes/no) as a predictor in our models. We found that a main-effects model had the best fit to the data—full

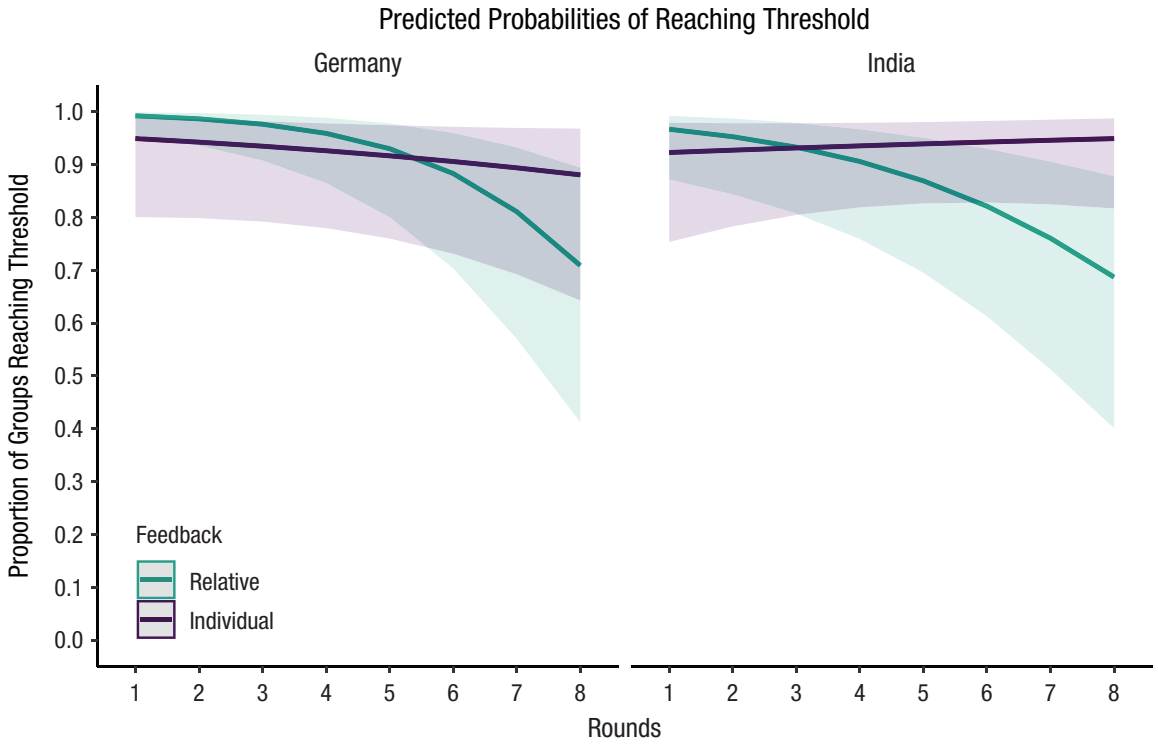

Fig. 4. Predicted rates of reaching the threshold based on the model with the best fit (two-way interaction model). Shaded areas show 95% confidence intervals.

versus null: $\chi^2(10) = 31.51$, $p < .001$; full versus two-way model: $\chi^2(1) = 0.84$, $p = .36$; two-way versus main-effects model: $\chi^2(3) = 2.51$, $p = .47$. This model revealed a significant effect of previously reaching the threshold, $Est = 1.16$, $SE = 0.45$, $\chi^2(1) = 6.56$, $p = .01$, with groups that reached the threshold in the previous round being more likely to reach it again in the current round. There were also significant effects of round, $Est = -0.23$, $SE = 0.08$, $\chi^2(1) = 8.91$, $p = .003$, and gender, $Est = 0.89$, $SE = 0.44$, $\chi^2(1) = 4.29$, $p = .038$. No other factors reached significance (see the Supplemental Material for details).

Overshooting the threshold. In further exploratory analyses, we investigated whether children overshot the threshold and thereby wasted water. Specifically, we defined overshooting as contributing more than 660 ml

Table 1. Outputs of the Two-Way Interaction Model Predicting Whether a Group Reached the Threshold

Parameter	Est	*SE*	χ^2	*df*	*p*
Intercept	6.75	1.83	n/a	n/a	n/a
Location[a]	−1.67	1.09	—	—	—
Feedback condition[b]	−2.35	1.06	—	—	—
Round	−0.57	0.13	—	—	—
Gender[c]	1.05	0.51	4.22	1	0.04
Average age of group	−0.17	0.21	0.68	1	0.41
Location × Feedback	1.03	1.03	1.00	1	0.32
Location × Round	0.20	0.14	2.12	1	0.15
Feedback × Round	0.43	0.14	10.71	1	0.001

Note: Estimates (Est) and standard errors (*SE*) are shown; *p* values for fixed effects are derived from likelihood ratio tests. *df* = degrees of freedom.
[a]Reference category: Germany. [b]Reference category: relative feedback.
[c]Reference category: male.

(110% of the threshold) in a round because there could have been some overflow of water from the connecting tubes even after the taps were closed. We found that a two-way model had the best fit to the data—full versus null: $\chi^2(9) = 138.50$, $p < .001$; full versus two-way: $\chi^2(1) = 0.33$, $p = .57$; full vs. main-effects model: $\chi^2(3) = 16.47$, $p = .001$. Overshooting happened mostly in the first rounds and dropped quickly (see the plots in the Supplemental Material), indicating a learning effect. It decreased more slowly in the individual-feedback condition than in the relative-feedback condition—Condition × Round: $Est = 0.46$, $SE = 0.16$, $\chi^2(1) = 8.81$, $p = .003$—and Indian children overshot for longer in the game than German children did—Location × Round: $Est = 0.47$, $SE = 0.18$, $\chi^2(1) = 7.78$, $p = .005$. Older children overshot less often than younger children, $Est = -0.84$, $SE = 0.27$, $\chi^2(1) = 9.98$, $p = .002$.

Water outcomes and variance in contributions. Water outcomes were similar in the two conditions in Germany (relative feedback: $M = 4{,}064$ ml, $SD = 488$ ml; individual feedback: $M = 3{,}963$, $SD = 555$ ml) and in India (relative feedback: $M = 3{,}842$ ml, $SD = 573$ ml; individual feedback: $M = 3{,}927$, $SD = 432$ ml). A full model with a Condition × Location interaction and main effects of gender and groups' average age (and group ID as a random intercept) did not differ significantly from a null model with only the random intercept, $\chi^2(5) = 6.87$, $p = .23$.

Exploratory analyses of within-group variation in contributions showed that a main-effects model had the best fit to the data—full versus null: $\chi^2(9) = 19.52$, $p = .021$; full versus two-way: $\chi^2(1) = 2.37$, $p = .12$; full versus main-effects model: $\chi^2(3) = 3.09$, $p = .38$—and revealed lower variance for groups in India than in Germany—$Est = -26.91$, $SE = 11.54$, $\chi^2(1) = 5.60$, $p = .018$. Plots of each group's contribution patterns can be found in the Supplemental Material.

Children's communication

We transcribed and coded children's utterances, resulting in a total of 14,374 codable utterances (89% of transcribed utterances). There were 16 main coding categories, which included, for example, verbal coordination strategies such as referring to giving water or to goals (e.g., "The seal needs to be in the water"), social comparisons (e.g., "I have a lot more than you"), fairness, promises, and threats. We also coded whether children used normative language (e.g., used sentences containing "should," "must," "have to"). Statistical analyses focused on the most frequent categories, and categories that were coded less than 1% of the time were only reported descriptively (see the Supplemental Material for an overview). Because there were more codable utterances in Germany (8,203 utterances, or 94% of transcribed utterances) than in India (6,171, or 83%), we conducted all analyses with proportions (i.e., utterances coded in a category in relation to the total number of codable utterances per participant group; see the Method section for details). The percentage of codable utterances was similar between conditions in Germany (individual-feedback condition: 3,586, or 92%; relative-feedback condition: 4,617, or 95%) and in India (individual: 3,149, or 83%; relative: 3,022, or 83%). We fitted GLMMs with the respective topic as a response variable, and feedback condition (individual vs. relative), location (Germany vs. India), coding category (where applicable), gender, and average age as fixed effects. Group number was entered as a random intercept (see the Method section). As before, we used model comparisons to determine the model with the best fit to the data.

On-topic utterances. Of the 14,374 codable utterances, 71.3% (10,242) were on topic (i.e., scored in at least one of the coding categories). There were more on-topic utterances in the relative-feedback condition than in the individual-feedback condition, both in Germany (relative: 72.9%, individual: 69.2%) and in India (relative: 74.7%; individual: 67.9%). Model comparisons showed that the full model had the best fit to the data—full versus null: $\chi^2(5) = 58.21$, $p < .001$; full versus main-effects model: $\chi^2(1) = 4.57$, $p = .033$. There was a significant interaction of Condition × Location, $Est = -0.16$, $SE = 0.08$, $\chi^2(1) = 4.57$, $p = .033$: children produced more on-topic utterances in the relative-feedback condition than the individual-feedback condition, with the difference between conditions being more pronounced in India compared to Germany. There was also a significant effect of gender, $Est = 0.09$, $SE = 0.04$, $\chi^2(1) = 5.93$, $p = .015$; male groups produced more on-topic utterances than female groups. No other factors were significant (see the Supplemental Material).

Verbal coordination. To get insight into children's verbal coordination, we analyzed references to giving water, to stopping to give water, and to goals. For each of these main categories, we coded subcategories that included the target of the utterance (e.g., themselves [self], another child [other], or the whole group [we]). For goals, we also distinguished whether participants referred to the group's common goal (e.g., "We have to win") or more generally to the goal (e.g., "If she [animal] is not inside, this means nobody gets a [water] bag"). References to giving water had the highest overall occurrence (34.1% of utterances). Children mentioned stopping to give water 10.9% of the time and mentioned goals 9.4% of the time, respectively.

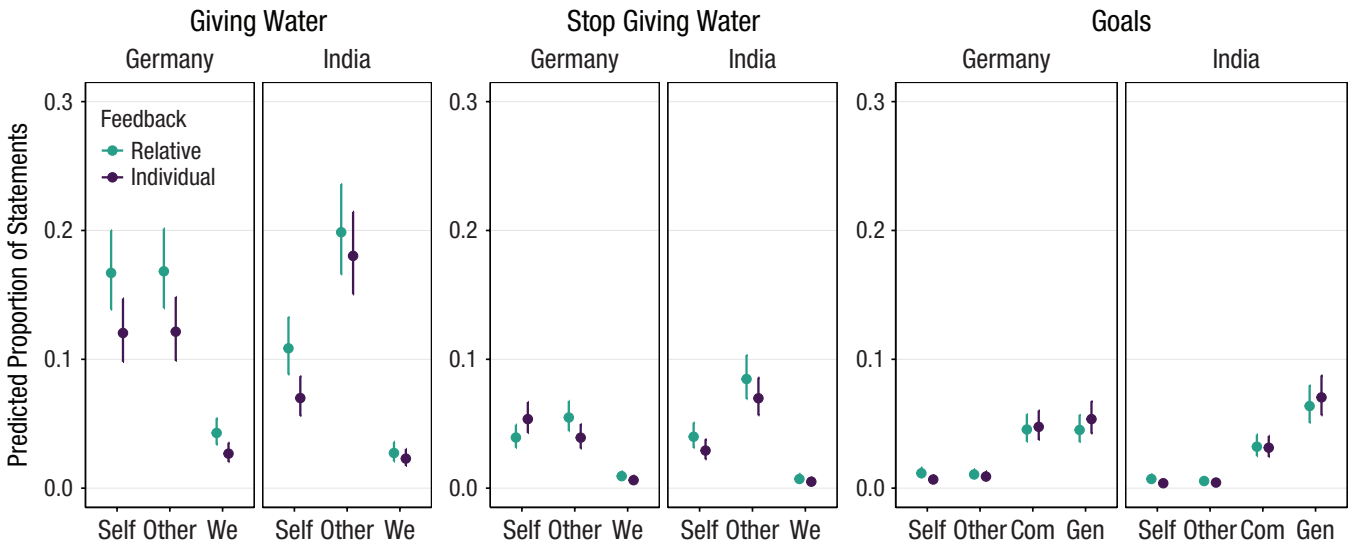

Fig. 5. Predicted proportions (with 95% confidence intervals) of groups' utterances based on the models with the best fit (giving: full model; stopping: full model; goals: two-way interaction model). "Self" indicates that children referred to themselves (e.g., "I give a lot," "I stop," "I have to win"). "Other" indicates that children referred to at least one other child in the group (e.g., "Fill man fill," "Close the tap," "You are not going to win this"). "We" indicates that children referred to the whole group (e.g., "Let us fill that," "We should slowly stop"). "Com" indicates that children referred to the group's common goals (e.g., "We have to win"), and "Gen" indicates that they referred to goals more broadly (e.g., "The seal has to be a bit in the water"). Note that utterances could be coded in more than one (sub)category. We used proportions (i.e., a group's utterances scored in a category in relation to the group's codable utterances), because there were overall more codable utterances in Germany (individual-feedback condition: 3,586, or 92% of total transcribed; relative-feedback condition: 4,617, or 95% of total transcribed) than in India (individual-feedback condition: 3,149, or 83% of total transcribed; relative feedback condition: 3,022, or 83% of total transcribed).

Concerning references to giving water, a full model (three-way interaction) had the best fit to the data—full versus null: $\chi^2(13) = 2{,}082.76$, $p < .001$; full versus two-way: $\chi^2(2) = 8.00$, $p = .018$. There was a significant three-way interaction of Location × Condition × Subcategory, $\chi^2(2) = 8.00$, $p = .018$. Children in Germany and India more frequently referred to giving in the relative-feedback condition than in the individual-feedback condition, with the effect being slightly attenuated in India for two of the three subcategories (referring to "others" or "we"; see Fig. 5). Moreover, Indian children were about twice as likely to refer to others than to themselves, whereas German children referred to themselves and others at similar rates. There was also a significant effect of age, $Est = 0.10$, $SE = 0.04$, $\chi^2(1) = 5.41$, $p = .020$, with older children referring to giving water more often than younger children. No other effects were significant.

Regarding stopping to give water, a full model (three-way interaction) again had the best fit to the data—full versus null: $\chi^2(13) = 830.37$, $p < .001$; full versus two-way: $\chi^2(2) = 12.65$, $p = .002$. There was a significant three-way interaction of Location × Condition × Subcategory, $\chi^2(2) = 12.65$, $p = .002$. Again, children in India referred more often to others than to themselves, but children in Germany referred to themselves and others at similar rates (see Fig. 5), with some variation across conditions in the two locations. No other effects were significant.

Analyzing children's references to goals, we found that the two-way interaction model had the best fit to the data—full versus null: $\chi^2(17) = 1{,}073.01$, $p < .001$; full versus two-way: $\chi^2(3) = 7.74$, $p = .052$; two-way versus main-effects model: $\chi^2(7) = 72.61$, $p < .001$)—and revealed a significant two-way interaction of Location × Subcategory, $\chi^2(3) = 56.56$, $p < .001$. Children in India were more likely to refer to general than to common goals, and children in Germany mentioned both about equally (see Fig. 5). There was also a significant Condition × Subcategory interaction, $\chi^2(3) = 13.45$, $p = .004$, but we are cautious about interpreting any patterns given the low occurrence of some goals (self, other) and the large overlap in 95% confidence intervals.

For exploratory analyses correlating "giving," "stopping," and goal-related utterances with behavior (e.g. reaching of threshold), see the Supplemental Material.

Social comparisons. Children referred to social comparisons (e.g., "Look how much water you have left. Look how little I have") 10.3% of the time. A main-effect model had the best fit to the data—full versus null: $\chi^2(5) = 159.37$, $p < .001$; full versus main-effects model: $\chi^2(1) = 0.28$, $p = .60$—and revealed a significant effect of condition, $Est = 0.24$, $SE = 0.06$, $\chi^2(1) = 18.91$, $p < .001$.

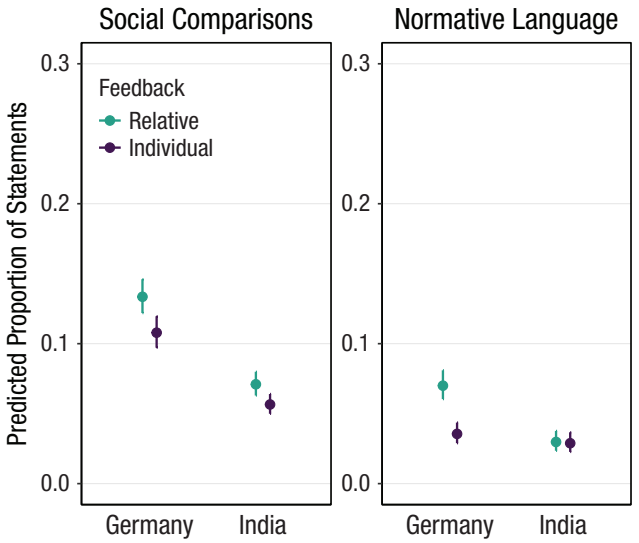

Fig. 6. Predicted proportions (with 95% confidence intervals) of groups' utterances based on the models with the best fit (social comparisons: main-effects model; normative language: full model). Note that utterances could be coded in more than one category.

Children in the relative-feedback condition referred to social comparisons more often than children in the individual-feedback condition (see Fig. 6). There was also a significant effect of location, $Est = 0.70$, $SE = 0.07$, $\chi^2(1) = 120.47$, $p < .001$: Children in Germany referenced social comparisons more often than children in India. Older children were more likely than younger children to utter social comparisons, $Est = 0.09$, $SE = 0.02$, $\chi^2(1) = 12.24$, $p < .001$, and male groups were more likely to do so than female groups, $Est = 0.14$, $SE = 0.06$, $\chi^2(1) = 5.85$, $p = .016$. For analyses exploring potential relations between social comparison utterances and behavior, see the Supplemental Material.

Normative language. Occasionally, children used normative language (3.7% of the time). A two-way interaction model had the best fit to the data—full versus null: $\chi^2(5) = 114.06$, $p < .001$; full versus two-way: $\chi^2(1) = 11.26$, $p = .001$—revealing a significant interaction of Location × Condition, $Est = 0.68$, $SE = 0.20$, $\chi^2(1) = 11.26$, $p < .001$. Children in Germany used normative language more frequently in the relative-feedback condition than the individual-feedback condition, but children in India rarely used normative language in either condition (see Fig. 6). Females were more likely than males to use normative language, $Est = 0.39$, $SE = 0.09$, $\chi^2(1) = 18.87$, $p < .001$.

Fairness, reciprocity, promises, threats, and other utterances. Children very rarely mentioned fairness (26 of 14,374, or 0.2% of codable utterances) or reciprocity (0.2%). Similarly, promises were rarely used (0.2%). Direct threats occurred almost never (0.04%), and few children mentioned tattling to the experimenter (0.3%). Children rarely mentioned cheating (0.7%). Few children complimented (0.2%) or insulted (0.4%) each other. For further details, see the Supplemental Material.

Discussion

We investigated the effect of outcome feedback on 6- to 10-year-old Indian and German children's cooperation in a threshold public-goods game. Cooperation remained robust for the duration of the game. Observing one's own outcomes relative to others' outcomes (relative feedback) led to a decrease in cooperation rates across rounds compared to seeing only one's own outcome (individual feedback). Nevertheless, the majority of groups reached the threshold even in the last round of the relative-feedback condition (India: 65%; Germany: 75%). This shows that, although children were sensitive to relative feedback (McClintock & McNeel, 1966; McClintock & Nuttin, 1969; Toda et al., 1978), thresholds can act as strong focal points for cooperation not only in adults (Barrett & Dannenberg, 2012; Cadsby & Maynes, 1999; Deutchman et al., 2022) but also in children.

We observed broadly similar cooperative behavior in Indian and German children and no effect of age. Moreover, exploratory analyses provided evidence for conditional cooperation in all age groups (Hermes et al., 2020). Previous cross-cultural research found societal variation in children's cooperative behavior in a collective risk-threshold game (Bowie et al., 2022) and in children's competitive behaviors (McClintock, 1974; McClintock & Nuttin, 1969; Toda et al., 1978), but evidence for age effects has been mixed (Angerer et al., 2016; Dutra et al., 2018; Harbaugh & Krause, 2000; Keil et al., 2017). The threshold likely helped children, regardless of age, to coordinate their actions and did so effectively in both locations.

We observed a small gender effect, with girls cooperating more than boys—similar to what had been found for German children in a common pool-water game (Koomen & Herrmann, 2018). However, there exists no consistent picture in the literature concerning gender differences in sharing and cooperation: Some studies have found no gender effects, and others have shown more prosocial behavior in girls or in boys, respectively (Angerer et al., 2015; Cárdenas et al., 2014; House et al., 2023; Sutter et al., 2019).

Analyses of utterances supported and enhanced the behavioral findings. Children referred to common and general goals (but rarely to their own or others' goals) in both conditions, indicating that groups were focused on reaching the threshold. However, children in Germany and India made social comparisons more frequently in the relative-feedback condition than in the individual-feedback condition—a clear indication that children paid more attention to group members'

contributions or outcomes. Children in both locations also produced more on-topic utterances and talked more about giving water in the relative-feedback condition (children in Germany also used more normative language). This shows that children increased their verbal effort to coordinate when seeing others' outcomes and flexibly responded to the different situational demands in the two conditions. Having the opportunity to communicate during the game likely contributed to maintaining high cooperation rates (Balliet, 2010; Sally, 1995).

Although we found broad similarities in communication patterns in response to situational demands (e.g., feedback) in Indian and German children, we also observed cultural nuances in children's communication styles. When talking about giving water or stopping their water contributions, Indian children referred more often to others than to themselves, whereas German children referred to themselves and others at similar rates. German children also verbalized social comparisons more frequently than Indian children. This aligns with previous studies on cultural models that found a stronger focus on others and interpersonal relationships in parent-child interactions in (urban) India than in (urban) Germany (Kärtner et al., 2016). Societal variation in communicative styles becomes relevant when collaborating and negotiating across cultures (Meyer, 2015)—a requirement for many pressing global issues.

It has been suggested that many real-world challenges, such as the climate emergency, can be conceptualized as threshold dilemmas (Barrett, 2007; Milinski et al., 2008). The potential for exploring threshold effects is by no means limited to public-goods games (Cadsby & Maynes, 1999; Dannenberg et al., 2015; Tavoni et al., 2011) but can also be extended, for example, to tragedy-of-the-commons games with harvesting thresholds (Ebersbach et al., 2019). Moreover, thresholds in our study were set externally, and future studies could have groups hold a binding vote on thresholds (Hauser et al., 2014).

We have shown that children's cooperative behavior remained robust to seeing others' outcomes in a threshold social dilemma. Clear and unambiguous thresholds (Dannenberg et al., 2015) can foster cooperation and provide a buffer against the most detrimental effects of social comparisons (Garcia et al., 2013).

Transparency

Action Editor: Leah Somerville
Editor: Patricia J. Bauer
Author Contributions

Patricia Kanngiesser: Conceptualization; Formal analysis; Funding acquisition; Methodology; Project administration; Supervision; Visualization; Writing – original draft; Writing – review & editing.

Jahnavi Sunderarajan: Project administration; Supervision; Writing – review & editing.

Sebastian Hafenbrädl: Conceptualization; Writing – review & editing.

Jan K. Woike: Conceptualization; Methodology; Writing – review & editing.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This study was supported by a Freigeist Fellowship from the Volkswagen Foundation to P. Kanngiesser (Grant No. 48952).

Open Practices

No aspects of the study were preregistered. The study materials, the data underlying the analyses reported in the article, and the analysis script are publicly available (https://osf.io/xzfce/). The study materials are also available in the Supplemental Material. Transcripts of children's conversations are available on request to the corresponding author. This article has received the badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iDs

Patricia Kanngiesser https://orcid.org/0000-0003-1068-3725
Sebastian Hafenbrädl https://orcid.org/0000-0002-5148-766X
Jan K. Woike https://orcid.org/0000-0002-6816-121X

Acknowledgments

We would like to thank Anna Puganovskaya, Iris Würbel, Elena Schmidt Velasco, Kathrin Kern, Süheyla Yilmaz, Isabel Lehmann, Käthi Wang, and Julia Ohlendorf for their help with this study. We would also like to thank the schools, parents, and children who participated in this study. We would like to acknowledge Maxi Friederike Woyke's help with reliability coding; Maxi sadly and unexpectedly passed away in 2021.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/09567976241267854

References

Alencar, A. I., de Oliveira Siquiera, J., & Yamamoto, M. E. (2008). Does group size matter? Cheating and cooperation in Brazilian school children. *Evolution and Human Behavior*, *29*(1), 42–48. https://doi.org/10.1016/j.evolhumbehav.2007.09.001

Angerer, S., Glätzle-Rützler, D., Lergetporer, P., & Sutter, M. (2015). Donations, risk attitudes and time preferences: A

study on altruism in primary school children. *Journal of Economic Behavior & Organization*, *115*, 67–74. https://doi.org/10.1016/j.jebo.2014.10.007

Angerer, S., Glätzle-Rützler, D., Lergetporer, P., & Sutter, M. (2016). Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city. *European Economic Review*, *90*, 254–264. https://doi.org/10.1016/j.euroecorev.2016.02.022

Balliet, D. (2010). Communication and cooperation in social dilemmas: A meta-analytic review. *Journal of Conflict Resolution*, *54*(1), 39–57. https://doi.org/10.1177/0022002709352443

Barrett, S. (2007). The smallpox eradication game. *Public Choice*, *130*(1–2), 179–207. https://doi.org/10.1007/s11127-006-9079-z

Barrett, S., & Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences, USA*, *109*(43), 17372–17376. https://doi.org/10.1073/pnas.1208417109

Barrett, S., & Dannenberg, A. (2014). Sensitivity of collective action to uncertainty about climate tipping points. *Nature Climate Change*, *4*(1), 36–39. https://doi.org/10.1038/nclimate2059

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Blake, P. R., Rand, D. G., Tingley, D., & Warneken, F. (2015). The shadow of the future promotes cooperation in a repeated prisoner's dilemma for children. *Scientific Reports*, *5*(1), Article 14559. https://doi.org/10.1038/srep14559

Boucher, E. (2024). Competition exams and topper praise in India: A media case study in educational (in)equality. In T. Saeed, R. Iyengar, M. A. Witenstein, & E. J. Byker (Eds.), *Exploring education and democratization in South Asia* (pp. 33–53). Springer.

Bowie, A., Zhou, W., Tan, J., White, P., Stoinski, T., Su, Y., & Hare, B. (2022). Motivating children's cooperation to conserve forests. *Conservation Biology*, *36*, Article e13922. https://doi.org/10.1111/cobi.13922

Butler, R. (1989). Mastery versus ability appraisal: A developmental study of children's observations of peers' work. *Child Development*, *60*, 1350–1361. https://doi.org/doi.org/10.2307/1130926

Cadsby, C. B., & Maynes, E. (1999). Voluntary provision of threshold public goods with continuous contributions: Experimental evidence. *Journal of Public Economics*, *71*(1), 53–73. https://doi.org/10.1016/s0047-2727(98)00049-8

Cárdenas, J.-C., Dreber, A., von Essen, E., & Ranehill, E. (2014). Gender and cooperation in children: Experiments in Colombia and Sweden. *PLOS ONE*, *9*(3), Article e90923. https://doi.org/10.1371/journal.pone.0090923

Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., & Tavoni, A. (2015). On the provision of public goods with probabilistic and ambiguous thresholds. *Environmental and Resource Economics*, *61*(3), 365–383. https://doi.org/10.1007/s10640-014-9796-6

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, *31*(1), 169–193. https://doi.org/10.1146/annurev.ps.31.020180.001125

Deutchman, P., Amir, D., Jordan, M. R., & McAuliffe, K. (2022). Common knowledge promotes cooperation in the threshold public goods game by reducing uncertainty. *Evolution and Human Behavior*, *43*(2), 155–167. https://doi.org/10.1016/j.evolhumbehav.2021.12.003

Dutra, N. B., Boccardi, N. C., Silva, P. R. R., de Oliveira Siqueira, J., Hattori, W. T., Yamamoto, M. E., & de Alencar, A. I. (2018). Adult criticism and vigilance diminish free riding by children in a social dilemma. *Journal of Experimental Child Psychology*, *167*, 1–9. https://doi.org/10.1016/j.jecp.2017.10.007

Ebersbach, M., Malkus, D., & Ernst, A. (2019). Factors that affect primary school children's sustainable behavior in a resource dilemma. *Journal of Experimental Child Psychology*, *184*, 18–33. https://doi.org/10.1016/j.jecp.2019.03.007

Gamer, M., Lemon, J., & Singh, I. F. P. (2019). *irr: Various coefficients of interrater reliability and agreement* (R Package Version 0.84.1). https://CRAN.R-project.org/package=irr

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on Psychological Science*, *8*(6), 634–650. https://doi.org/10.1177/1745691613504114

Grueneisen, S., & Tomasello, M. (2017). Children coordinate in a recurrent social dilemma by taking turns and along dominance asymmetries. *Developmental Psychology*, *53*(2), 265–273. https://doi.org/10.1037/dev0000236

Grueneisen, S., & Tomasello, M. (2019). Children use rules to coordinate in a social dilemma. *Journal of Experimental Child Psychology*, *179*, 362–374. https://doi.org/10.1016/j.jecp.2018.11.001

Harbaugh, W. T., & Krause, K. (2000). Children's altruism in public good and dictator experiments. *Economic Inquiry*, *38*(1), 95–109. https://doi.org/10.1111/j.1465-7295.2000.tb00006.x

Hauser, O. P., Rand, D. G., Peysakhovich, A., & Nowak, M. A. (2014). Cooperating with the future. *Nature*, *511*(7508), 220–223. https://doi.org/10.1038/nature13530

Hermes, H., Hett, F., Mechtel, M., Schmidt, F., Schunk, D., & Wagner, V. (2020). Do children cooperate conditionally? Adapting the strategy method for first-graders. *Journal of Economic Behavior & Organization*, *179*, 638–652. https://doi.org/10.1016/j.jebo.2018.12.032

House, B. R., Kanngiesser, P., Barrett, H. C., Broesch, T., Cebioglu, S., Crittenden, A. N., Erut, A., Lew-Levy, S., Sebastian-Enesco, C., Smith, A. M., Yilmaz, S., & Silk, J. B. (2020). Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nature Human Behaviour*, *4*(1), 36–44. https://doi.org/10.1038/s41562-019-0734-z

House, B. R., Silk, J. B., Henrich, J., Barrett, H. C., Scelza, B. A., Boyette, A. H., Hewlett, B. S., McElreath, R., & Laurence, S. (2013). Ontogeny of prosocial behavior across diverse societies. *Proceedings of the National*

Academy of Sciences, USA, *110*(36), 14586–14591. https://doi.org/10.1073/pnas.1221217110

House, B. R., Silk, J. B., & McAuliffe, K. (2023). No strong evidence for universal gender differences in the development of cooperative behaviour across societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1868), Article 20210439. https://doi.org/10.1098/rstb.2021.0439

Houser, D., Montinari, N., & Piovesan, M. (2012). Private and public decisions in social dilemmas: Evidence from children's behavior. *PLOS ONE*, *7*(8), Article e41568. https://doi.org/10.1371/journal.pone.0041568

Kanngiesser, P., Sunderarajan, J., & Woike, J. K. (2023). Cheating and the effect of promises in Indian and German children. *Child Development*, *95*, 16–23. https://doi.org/10.1111/cdev.13952

Kärtner, J., Crafa, D., Chaudhary, N., & Keller, H. (2016). Reactions to receiving a gift—Maternal scaffolding and cultural learning in Berlin and Delhi. *Child Development*, *87*(3), 712–722. https://doi.org/10.1111/cdev.12525

Keil, J., Michel, A., Sticca, F., Leipold, K., Klein, A. M., Sierau, S., von Klitzing, K., & White, L. O. (2017). The Pizzagame: A virtual public goods game to assess cooperative behavior in children and adolescents. *Behavior Research Methods*, *49*(4), 1432–1443. https://doi.org/10.3758/s13428-016-0799-9

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology*, *24*(1), 183–214. https://doi.org/10.1146/annurev.soc.24.1.183

Koomen, R., & Herrmann, E. (2018). An investigation of children's strategies for overcoming the tragedy of the commons. *Nature Human Behaviour*, *2*(5), 348–355. https://doi.org/10.1038/s41562-018-0327-2

Lee, H., Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W. W. L., Connors, S. L., Denton, F., Diongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., . . . Zommers, Z. (2023). *Synthesis report of the IPCC Sixth Assessment Report (AR6): Summary for policymakers*. https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_SPM.pdf

Lüdecke, D. (2023). *Sjplot: Data visualization for statistics in social science* (R Package Version 2.8.14). https://CRAN.R-project.org/package=sjPlot

McClintock, C. G. (1974). Development of social motives in Anglo-American and Mexican-American children. *Journal of Personality and Social Psychology*, *29*(3), 348–354. https://doi.org/10.1037/h0036019

McClintock, C. G., & McNeel, S. P. (1966). Reward and score feedback as determinants of cooperative and competitive game behavior. *Journal of Personality and Social Psychology*, *4*(6), 606–613. https://doi.org/10.1037/h0023986

McClintock, C. G., & Moskowitz, J. M. (1976). Children's preferences for individualistic, cooperative, and competitive outcomes. *Journal of Personality and Social Psychology*, *34*(4), 543–555. https://doi.org/10.1037/0022-3514.34.4.543

McClintock, C. G., Moskowitz, J. M., & McClintock, E. (1977). Variations in preferences for individualistic, competitive, and cooperative outcomes as a function of age, game class, and task in nursery school children. *Child Development*, *48*(3), 1080–1085. https://doi.org/10.2307/1128365

McClintock, C. G., & Nuttin, J. M. (1969). Development of competitive game behavior in children across two cultures. *Journal of Experimental Social Psychology*, *5*(2), 203–218. https://doi.org/10.1016/0022-1031(69)90047-X

Messick, D. M., & McClintock, C. G. (1968). Motivational bases of choice in experimental games. *Journal of Experimental Social Psychology*, *4*(1), 1–25. https://doi.org/10.1016/0022-1031(68)90046-2

Meyer, E. (2015). Getting to sí, ja, oui, hai, and da. *Harvard Business Review*, *93*(12), 74–80.

Milinski, M., Sommerfeld, R. D., Krambeck, H.-J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences, USA*, *105*(7), 2291–2294. https://doi.org/10.1073/pnas.0709546105

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review*, *86*(2), 404–417. https://doi.org/10.2307/1964229

Prétôt, L., & McAuliffe, K. (2020). Does nonbinding commitment promote children's cooperation in a social dilemma? *Journal of Experimental Child Psychology*, *200*, Article 104947. https://doi.org/10.1016/j.jecp.2020.104947

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Sally, D. (1995). Conversation and cooperation in social dilemmas. *Rationality and Society*, *7*(1), 58–92. https://doi.org/10.1177/1043463195007001004

Steinbeis, N., & Singer, T. (2013). The effects of social comparison on social emotions and behavior during childhood: The ontogeny of envy and Schadenfreude predicts developmental changes in equity-related decisions. *Journal of Experimental Child Psychology*, *115*(1), 198–209. https://doi.org/10.1016/j.jecp.2012.11.009

Sutter, M., Zoller, C., & Glätzle-Rützler, D. (2019). Economic behavior of children and adolescents – a first survey of experimental economics results. *European Economic Review*, *111*, 98–121. https://doi.org/10.1016/j.euroecorev.2018.09.004

Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences, USA*, *108*(29), 11825–11829. https://doi.org/10.1073/pnas.1102493108

Toda, M., Shinotsuka, H., McClintock, C. G., & Stech, F. J. (1978). Development of competitive behavior as a function of culture, age, and social comparison. *Journal of Personality and Social Psychology*, *36*(8), 825–839. https://doi.org/10.1037/0022-3514.36.8.825

Tomasello, M. (2022). The coordination of attention and action in great apes and humans. *Philosophical Transactions of*

the Royal Society B, *377*(1859), Article 20210093. https://doi.org/10.1098/rstb.2021.0093

van de Kragt, A. J. C., Orbell, J. M., & Dawes, R. M. (1983). The minimal contributing set as a solution to public goods problems. *American Political Science Review*, *77*(1), 112–122. https://doi.org/10.2307/1956014

van Lange, P. A., Joireman, J., Parks, C. D., & van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, *120*(2), 125–141. https://doi.org/10.1016/j.obhdp.2012.11.003

Vogelsang, M., Jensen, K., Kirschner, S., Tennie, C., & Tomasello, M. (2014). Preschoolers are sensitive to free riding in a public goods game. *Frontiers in Psychology*, *5*, Article 729. https://doi.org/10.3389/fpsyg.2014.00729

Woike, J. K., & Hafenbrädl, S. (2020). Rivals without a cause? Relative performance feedback creates destructive competition despite aligned incentives. *Journal of Behavioral Decision Making*, *33*(4), 523–537. https://doi.org/10.1002/bdm.2162