ARTICLES
https://doi.org/10.1038/s41562-018-0327-2

nature human behaviour

Corrected: Publisher Correction

# An investigation of children's strategies for overcoming the tragedy of the commons

Rebecca Koomen* and Esther Herrmann

**Common-pool resource (CPR) dilemmas are pervasive challenges to overcome. We presented six-year-old children with an experimental CPR paradigm involving a renewable water resource, which children could collect to win individual rewards. To maximize water collection, children had to wait for water to accumulate, without collapsing the resource. We explore the social strategies children used to overcome the dilemma together. Like adults, six-year-old children were challenged by the dilemma: resource sustaining was more successful in a parallel condition in which children worked independently compared with the collective CPR condition. However, children were capable of collectively preventing resource collapse by spontaneously generating inclusive rules, equally distributing the rewards and distracting one another from the delay-of-gratification task. Children also learned to sustain the resource longer in repeated interactions with the same partner. Already by the age of six, children are capable of CPR social strategies resembling those of adults.**

Throughout history, humans have had to overcome resource dilemmas to sustain essential renewable resources. When naturally renewable resources are openly accessible and subtractable—meaning any amount taken by one individual subtracts from the total pool available to others at that time—a social dilemma can emerge in which all actors must maintain suboptimal harvests while enforcing the same on potential free riders[1]. Overharvesting a common-pool resource (CPR) can lead to the collapse of its renewal potential, resulting in resource depletion. Crucially, because the costs of overharvesting are shared by the group, incentives to free ride by taking too much at any time have lead economists to caution that CPR dilemmas, quite rationally, face the risk of resource collapse, as in the tragedy of the commons[2]. Thus, despite our pervasive cooperative skills[3], humans face significant collective action challenges in the face of CPR dilemmas.

Decades of field studies complemented by experiments in the laboratory have shown that humans are capable of coordinating action to overcome the tragedy of the commons when certain conditions are met (for an overview, see ref. [1]). Commonalities of successfully managed CPRs in the real world indicate the importance of rules regarding resource allocation, as well as monitoring and sanctioning mechanisms. Importantly, the social and ecological setting must allow for local, internal rules and sanctioning as opposed to exclusive external enforcement from, for example, a national government. The environment must also provide clear information on the state of the resource (that is, the renewal rate, quantity available and collapse risk). When groups have the authority to decide on their own rules and when they can communicate face-to-face, they have a higher likelihood of establishing trust, in-group identity and normative moral obligations to achieve group success[1]. An example of such locally resolved CPR dilemmas can be found in lobster-fishing villages in Maine. Many communities have developed local systems of mapping fishing territories throughout their accessible waters[4], which welcome or prohibit certain fishers based on social dominance, adherence to valued community norms signalling trustworthiness, and merit-based prestige. Free riders caught fishing in someone else's territory are punished by community members[4].

Inequality can cause cooperation to break down in social dilemmas over resources[1,5,6]. Experimental paradigms with adult humans highlight the importance of direct communication and the internal development of rules for modifying behaviour under conditions of inequality[7]. For example, when people are given unequal endowments at the beginning of an experimental multi-player CPR dilemma, they have a strong tendency to overexploit the resource, leading to its demise. However, when face-to-face communication is introduced between rounds, groups become much more successful at sustaining the CPR together, coordinating behaviour through proportional allocation rules that equalize net payoffs[8]. Likewise, if communication itself is unequal, as in experiments in which only a subset of subjects in an experimental CPR are given the chance to communicate with one another, success is not as high[9]. These findings were extended by allowing anonymous participants to communicate freely via computer messages[10], and it was found that group payoff from the CPR was higher when contributions to group discussions were equally distributed among participants.

Communication among actors in a social dilemma also helps promote important norms of reciprocity, fairness[11] and group identity[12]. Norms comprise beliefs and preferences about how others will act and internalized motivations to align one's behaviour with that of others[13]. Indeed, public goods experiments repeatedly indicate that willingness to contribute to a public good is linked to expectations of the behaviour of others involved (for example, refs [14–16]). Normative mechanisms help align cooperative behaviour in CPR dilemmas in a number of ways, such as activating in-group comparisons[17], highlighting fairness of allocations[18], enforcing religious or spiritual taboos to limit resource consumption[19], regulating behaviour as a substitute for formal institutions or incomplete contracts[20], and promoting collective action[21]. Integrating two decades of empirical and theoretical CPR research, ref. [22] identifies shared group norms and fairness of resource allocation to be two of the most significant predictors of successful CPR management.

To date, no experimental CPR paradigms have explored the behaviour of subjects other than adult humans. As such, little is known about children's abilities to navigate resource dilemmas and the development of the social and cognitive skills necessary

Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. *e-mail: becky.koomen@gmail.com

to overcome the tragedy of the commons. Furthermore, studying children in a resource dilemma allows for unique insight into how cognitive development influences cooperative behaviour in a social dilemma with real-world relevance. Ref. [23] identified five cognitive prerequisites for children's abilities in economic games: (1) beliefs about others and theory of mind (these emerge around four to five years of age[24–26]); (2) differentiating between intentional and unintentional actions of agents (this emerges by the age of four[27,28]); (3) connecting comprehension of others' beliefs and one's own decisions (for example, through cheating or deceiving; this is already present in four-year-olds[29]); (4) coordination skills (these are present in five-year-olds through group conformity mechanisms[30] and in six-year-olds as first- and second-order false belief understanding[31]); and (5) distributive justice and norms for fair resource allocation. Sensitivity to equal distributions already appears in infants[32] and by the age of three, children start to expect fair distributions for themselves and others, incorporating more complex principles of fairness relative to the situation (that is, norms, merit, need, and so on) by the age of five (ref. [33]).

Taken together, these developmental findings indicate that by five to six years of age, children possess the basic cognitive building blocks necessary to navigate social dilemmas. Although relatively little work has been done with children in the context of economic or social dilemmas, some studies indicate support for this conclusion. For example, ref. [34] showed that five-year-olds spontaneously generated rules to overcome a chicken-game dilemma, in which one child had to volunteer for a smaller reward to ensure that both children received a reward. Five-year-olds were also tested in groups of four in a mini-public-goods game (PGG) in which they could anonymously allocate two gumballs to themselves or four gumballs to a public pot. Individual payoff was highest when three children chose the public option and the focal child chose the private option[35] (see also ref. [36]). Initial contributions to the public good were relatively high but dropped in subsequent rounds due to individual free riding. This pattern indicates conditional cooperation, which adults show in a PGG[37] and five-and-a-half-year-old children show in a repeated prosocial choice task[38]. Furthermore, five-year-olds overcame a snowdrift conflict with unequal reward distribution in which participants had to wait for their partner to act in order to obtain the larger reward, but if no child acted all rewards would be lost[39]. Most pairs of children succeeded with a strategy that involved both partners acting, using imperative communication to coerce cooperation when this was not effective.

Like other social dilemmas, CPR dilemmas involve a conflict between individual and group benefits; however, they also involve an additional temporal component. Individual rewards are available immediately but group rewards are delayed, and the size of this delay influences the difficulty of cooperating to sustain a CPR[40]. Children's delay-of-gratification skills have been extensively studied using the popular marshmallow test developed by Mischel and colleagues (for a review, see ref. [41]). Children in this paradigm, often as young as four, are able to self-impose delay maintenance for up to 15 min when faced with one sweet reward and the possibility of a second if they do not eat the first by the end of a waiting period. Children wait much longer when they are able to divert their attention away from the delay task, using creative self-distraction strategies such as talking to themselves, singing and inventing spontaneous games[42]. It has been shown that six-year-olds—but not four-year-olds—are aware of visual, cognitive and motor distraction strategies to aid in delay maintenance[43]. Furthermore, children's performance in a delay-of-gratification task can be influenced by the presence of groupmates[44] and conformity to group choices[45].

Despite the vast literature on children's delay-of-gratification skills, the question of how these skills are affected by social interdependence—as in a CPR dilemma in which participants have to collectively forego immediate gains to avoid resource collapse—remains a relatively unexplored research topic. As such, this study aims to gain an understanding of the role of social interdependence on a collective delay-of-gratification task in children, as well as explore the social strategies children use to overcome CPR dilemmas. To do this, we presented dyads of six-year-old children with a renewing, open access, subtractable magic water resource (water with blue food colouring). Children could collect the magic water to assemble eggs for individual rewards, yet taking too much magic water at any time could collapse the resource, preventing any further accumulation of water from its renewal source (see Figs. 1 and 2 for an illustration of the experimental setup). The renewal rate and amount of available magic water were made visually salient with a floating red cork and a red threshold line. At no point in the experiment were children given an indication of pre-existing rules to play the game or how the adult experimenters expected them to play, to allow for dyads to develop their own rules of use. We compared dyads' behaviour in a collective condition, in which both children accessed magic water from the same source and were thus engaged in a CPR dilemma, with dyads in a parallel condition, in which each child had their own independent source of renewing magic water. The collective condition represented a CPR dilemma, whereas the parallel condition presented children with an independent dilemma in the same social and physical context as the collective condition, but in the absence of social interdependence. We examined differences in success between the two conditions and the effect of verbal contribution equality, reward equality, collectively inclusive verbal strategies and verbal self-distraction between partners on success in the collective condition (see Table 1 for verbal strategy level classification scale).

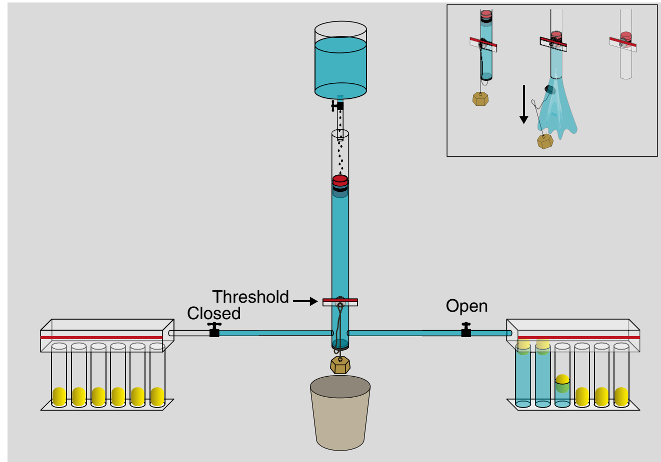


**Fig. 1 | Renewing magic water resource.** Taps could be opened or closed allowing children to control the flow of water, which filled each successive parallel cylinder. Empty yellow eggs sat at the bottom of each cylinder and floated to the top with the water level. The red cork inside the central cylinder floated up and down with the water level, rising when both taps were closed due to the inflow of water from the top source dripping pump, and lowering when the taps were open. If the red cork reached the threshold line, an automatic magnet mechanism caused a plug to open and all the water dropped into the bucket out of reach. All the water that continued to drip from the pump at the top also flowed straight into the bucket thereafter. Inset: enlarged collapse action mechanism.

## Results

An overview of dyadic success at the trial level in the two conditions is shown in Table 2.

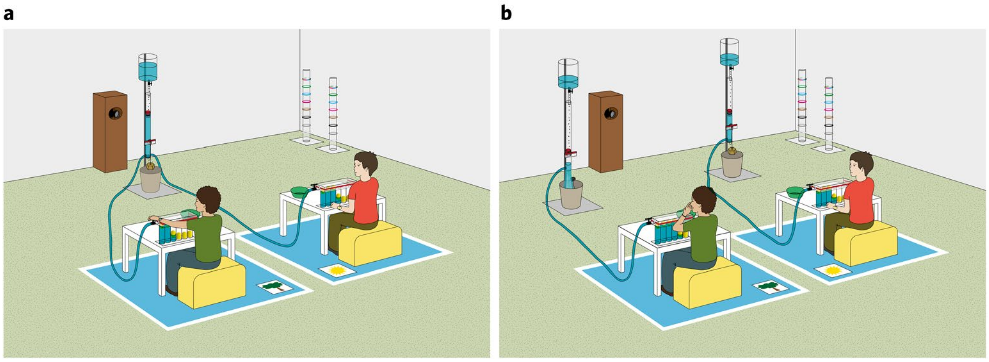


**Fig. 2 | Experimental setup. a**, Setup in the collective condition. **b**, Setup in the parallel condition, in which the child on the left has just collapsed his apparatus.

The first model, analysing the dependent measure of dyadic success, revealed that the interaction between condition and trial number was not significant (chi-squared test, $\chi^2 = 1.18$, d.f. = 1, $P = 0.277$) but the main effects of condition ($\chi^2 = 20.55$, d.f. = 1, $P < 0.001$) and trial number ($\chi^2 = 5.99$, d.f. = 1, $P = 0.014$) both had a significant effect on dyadic success. Dyads were more successful in the parallel condition than the collective condition, and dyads in both conditions became more successful across the three trials (Fig. 3). Female dyads also achieved higher success in both conditions than male dyads ($\chi^2 = 5.08$, d.f. = 1, $P = 0.024$); however, as this was not a test predictor we will not endeavour to interpret these results further.

Neither verbal equality ($\chi^2 = 0.3$, d.f. = 3, $P = 0.585$), nor proportional self-distraction ($\chi^2 = 1.02$, d.f. = 1, $P = 0.314$) varied significantly between conditions. However, a subsequent evaluation of the variation of verbal equality in the collective condition revealed that out of 56 trials, 12 had a verbal equality score of 0 and the remaining 44 had verbal equality scores ranging from 60–100%, indicating generally that partners' relative contributions to discussions were highly equal. Likewise, in the parallel condition, of 66 trials only 8 had a verbal equality score of 0, only 4 ranged from 1–59% and the remaining 54 ranged from 65–100%. All 20 trials with complete inequality of verbal contributions produced only one or two total dyadic utterances before collapse, indicating that complete inequality was only observed in trials in which one child did not speak at all and their partner communicated only briefly. We ran the same verbal equality model without these 20 trials ($n = 102$) and the results did not change ($\chi^2 = 0.12$, d.f. = 1, $P = 0.73$). Likewise, when analysing the likelihood of reaching terminal strategy (level 4), no significant difference between conditions was observed ($\chi^2 = 0.12$, d.f. = 1, $P = 0.73$). Therefore, for all five measures tested between conditions—success, the learning effect across the three trials, verbal equality, proportional self-distraction and the likelihood of reaching terminal verbal strategy—only success differed between conditions. The four behavioural variables—the learning effect, verbal equality, proportional self-distraction and the likelihood of reaching terminal verbal strategy—did not vary as a function of the outcome interdependence in the collective condition and the independence of the parallel condition.

Next, we analysed dyadic success in the collective condition only. Egg equality ($\chi^2 = 4.04$, d.f. = 1, $P = 0.044$) and verbal strategy level ($\chi^2 = 33.85$, d.f. = 1, $P < 0.001$) significantly predicted success in the collective condition. The closer dyads were to an equal 50:50 split of dyadic eggs (that is, when partners got approximately the same number of eggs as each other), the better they did as a dyad at sustaining and accumulating the water resource. Likewise, dyads were more successful during trials in which they reached more collectively inclusive verbal strategies than trials in which only independent or imperative verbal strategies were used. See Figs. 4 and 5 for egg equality and verbal strategy results, respectively.

The second model to analyse dyadic success in the collective condition only revealed that proportional self-distraction ($\chi^2 = 12.19$, d.f. = 1, $P < 0.001$), but not verbal equality ($\chi^2 = 0.4$, d.f. = 1, $P = 0.526$), significantly predicted dyadic success in the collective condition. Dyads were more successful at sustaining and accumulating the water resource when they spoke proportionally more about non-game-related topics during play, but the equality of verbal contributions between partners did not affect success. See Fig. 6 for the effect of proportional self-distraction on dyadic success.

**Table 1 | Verbal strategy-level descriptions and examples, increasing in collective inclusivity from 1 to 4**

| Level | Description | Example |
|---|---|---|
| 1 | No strategy | 'I have one egg up.' |
| 2 | Imperative or individual strategy without negotiation | 'Hurry! Close your tap!' |
| 3 | Imperative or individual strategy with agreement/rejection | Child 1: 'I am taking more water now' |
| | Collective strategy without agreement/rejection | Child 2: 'No leave it, otherwise you'll lose!' |
| 4 | Collective strategy with agreement/rejection | Child 1: 'Now we can both take a small drop' |
| | | Child 2: 'Yes now, look!' |

## Discussion

Although not always successful, dyads of six-year-old children are able to collectively sustain a CPR dilemma. Approximately 40% of dyads (11 out of 29) in the collective condition were successful in overcoming the CPR dilemma (that is, collected the maximum

**Table 2 | Distribution of trials according to corrected success per condition**

| Percent of available eggs collected corrected for minimum success | 0–25% | 26–50% | 51–75% | 76–100% |
|---|---|---|---|---|
| Collective trials ($n = 87$) | 56 | 6 | 3 | 22 |
| Parallel trials ($n = 75$) | 9 | 8 | 3 | 55 |

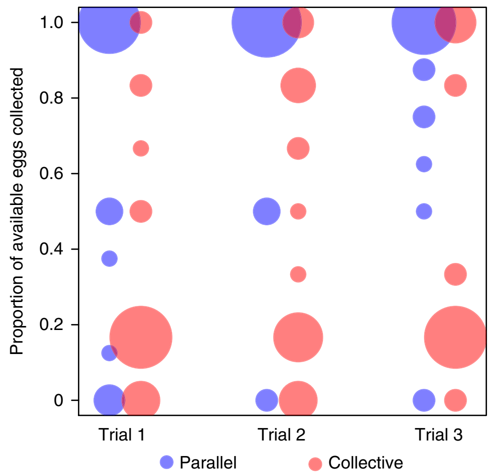


**Fig. 3 | Dyadic success across trials for each condition.** Dyadic success (*y* axis) is measured by the proportion of available eggs collected by dyads, corrected for differential egg potentials in the two conditions. Parallel condition dyads (blue) were significantly more successful than collective condition dyads (red), and all dyads increased in success across the three trials. The area of the circles (red and blue) is proportional to the frequency of data points at that location; larger circles represent a higher occurrence of data points ($n=162$ trials; 87 in the collective condition, 75 in the parallel condition).

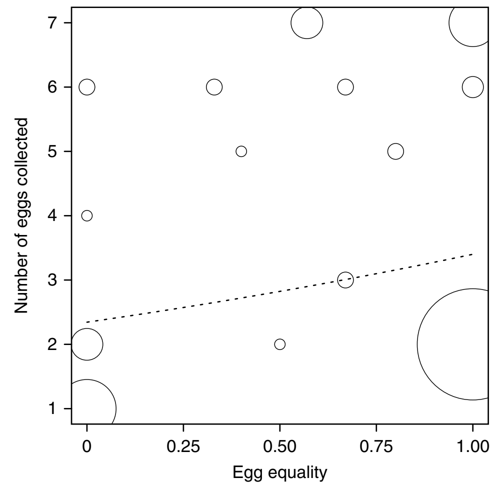


**Fig. 4 | Collective condition success and egg equality.** Success plotted as a function of egg equality. The dotted line represents the prediction of the model, increasing in success from complete inequality (0) to equality (1). The area of circles represents the relative frequency of data points at that position ($n=72$ collective condition trials).

number of eggs (6–7) in at least 1 of 3 trials). In comparison, 84% of dyads (21 out of 25) were successful in sustaining their independent water systems in the parallel condition without the competition and interdependence of the CPR dilemma. Because dyads in the collective condition were significantly less successful than dyads in the parallel condition, and because more than half of the collective condition trials were not sustained beyond the minimal success level of 1–2 eggs (64% of trials), we can conclude that CPR dilemmas are indeed challenging for six-year-old children, as is observed in studies with adults[7]. However, avoiding the tragedy of the commons is not impossible for six-year-olds.

Children in both conditions showed comparable numbers of self-distracting verbal exchanges, were equally likely to reach the highest level of collectively inclusive verbal strategy and did not differ in the degree of verbal equality between partners. These findings demonstrate that the social setting of the resource accumulation task—irrespective of payoff interdependence—elicited similar social responses in six-year-old children. Indeed, experiments with adults show that verbal strategizing, including verbal pleas to incite cooperative and norm-abiding behaviour, is prevalent in adults even when payoff to the speaker is not affected[46], such as in our parallel condition. Both conditions required children to maintain delay-of-gratification to maximize payoffs. The finding that self-distracting verbal responses did not differ between conditions is therefore consistent with evidence from delay-of-gratification studies showing that decreased attention paid to the object of reward—in this case the egg collecting game—aids in children's delay maintenance[41,47]. Such so-called 'attention deployment' is associated with increased delay-of-gratification success in 2-year-olds[48], 3–5-year-olds[42] and 6–12-year-olds[49]. However, the self-distraction measure in the present study was inherently social compared with previous studies, indicating that children may be able to use peers as distraction agents irrespective of outcome interdependence or competition over the resource.

In both conditions, children showed a learning effect with experience across the three trials, demonstrating that even with the competition and interdependence of the collective condition they were able to implement strategies to overcome the dilemma with increasing success. This finding is noteworthy when compared with findings from repeated-interaction PGGs with children, which showed a pattern of first increasing, plateauing, then declining contributions[35,36]. Moreover, repeated interactions in CPR-framed dilemmas with adults also show a typical pattern of increasing depletion rates unless institutional and communicative affordances are provided to structure allocation agreements and punish free riders[7,50]. This difference is potentially attributable to the fact that children in the present study sat in close physical proximity to one another and made resource extraction decisions that were visible to their partners at all times. This is in contrast with conventional PGG and CPR methodologies in which decisions are often anonymous at the time they are made.

Investigating the nature of dyadic success in the collective condition, we found that higher equality of payoff between partners significantly predicted success, and higher proportional self-distraction during trials as well as higher collective inclusivity of verbally discussed game strategies significantly improved success in the dilemma. These results suggest that by the age of six, children are capable of spontaneously implementing social strategies for overcoming CPR dilemmas similar to those shown in adults[7]. Although equality of verbal contributions by partners did not significantly predict success, the distribution of equality measures for this behavioural variable was heavily skewed towards equality in both conditions—children naturally tended towards symmetrical conversations. The lack of a significant effect for verbal contribution equality here can therefore potentially be explained by the fact that children already had a strong tendency to contribute to discussions once a discussion was initiated. Additionally, heterogeneity of discussion contributions among multiple actors in a group dilemma, as was previously measured in adult experiments[10], is likely to be more salient to individual actors than between two individuals in a dyadic dilemma.

In analysing the content of the verbal dialogue in the collective condition, nearly half of all dyads (12 out of 23) spontaneously generated discussions reaching the highest of our verbal strategy levels: collective strategy with agreement or rejection. Children in these trials ($n=20$) were significantly more likely to maximize their dyadic payoff. We found very little evidence in children's utterances of normative vocabulary. Perhaps because the task was novel to children, existing norms of resource allocation fairness

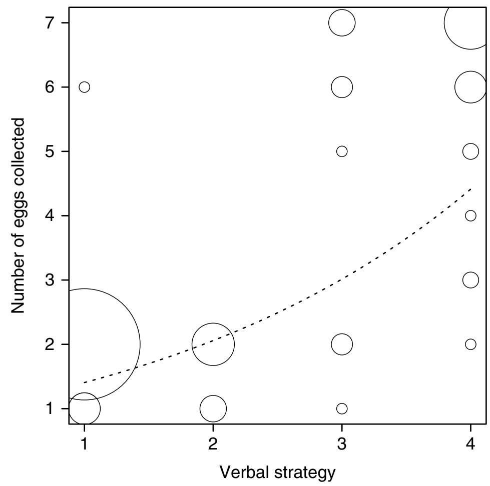


**Fig. 5 | Collective condition success and verbal strategy.** Success plotted as a function of verbal strategy levels 1–4. The dotted line represents the prediction of the model, increasing in success from no verbal strategy (level 1) to collective verbal strategy (level 4). The area of circles represents the relative frequency of data points at that position (*n* = 72 collective condition trials).

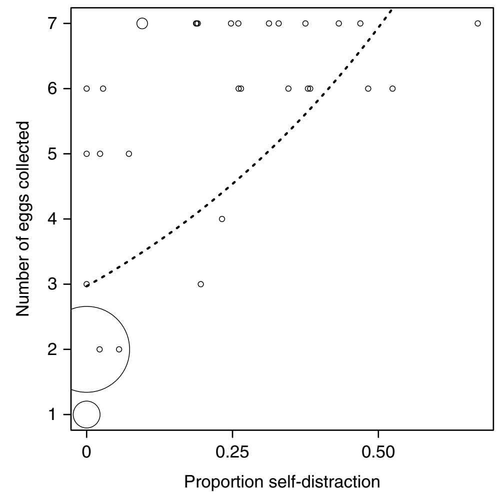


**Fig. 6 | Collective condition success and proportional self-distraction.** Success plotted as a function of proportional self-distraction. The dotted line represents the prediction of the model, increasing in success from zero self-distraction to approximately 50%. The area of circles represents the relative frequency of data points at that position (*n* = 51 collective condition trials excluding trials with zero categorizable utterances).

did not apply to their discussions in forming their own rules. A similar finding was reported by ref. [34], in which children had to coordinate decisions in a repeated chicken game. Children negotiated turn-taking strategies to improve dyadic payoff without applying normative language to achieve a fairer result. This led us to distinguish instead between spontaneously generated game strategies involving individual or imperative verbal rules that guided the behaviour of only one of the two children and negotiated, collectively inclusive strategies (including turn-taking and synchronicity) that specifically applied expectations to both children's behaviour. A closer post-hoc look at the types of collective strategies proposed indeed revealed two general distinctions in strategy type: synchronicity and turn-taking. Of the 55 collective strategies proposed within these 20 trials, 40 involved proposals relating either to waiting time, when water could or could not be accessed, or the amount of water taken by both individuals at the same time and to an equal degree. In contrast, only 15 involved explicit turn-taking proposals regarding directly asynchronous behaviour such that one individual proposed the timing or amount of water they or their partner could take while either they or their partner waited to take a turn. While we did not identify explicitly normative language in the game setting, we show that children's behaviour in the collective condition was consistent with what would be expected by a 50:50 resource-distribution pattern, in line with normative rules of fairness. Moreover, because normative understanding in children is most commonly measured through reactions to norm violations[51], we subsequently also examined policing behaviours. Although not explicitly normative, policing did occur in the collective condition transcript: a post-hoc search for 'stop!' and 'close your tap!' revealed that children protested in a little over 5% of total collective condition utterances (232 out of 4,329).

Like inclusive verbal strategies, equal payoffs between partners were also likely to improve dyadic payoff in the collective dilemma. As with inclusive verbal strategies, equally allocating the available rewards by collectively sustaining a CPR appears to be moderated at least in part by norms of fairness with respect to resource distribution. Equality (or inequality) in experiments with adults is commonly operationalized as differences in initial endowments; that is, funds necessary to access the resource (for example, ref. [8]). However, according to ref. [52], heterogeneity (that is, inequality) in almost all forms measured has a negative impact on cooperation in a CPR dilemma through processes affecting either incentives to cooperate or the effect that inequality has on social norms promoting cooperation and sanctioning enforcement. To learn more about how children apply norms of fairness or distributive justice to CPR strategies without directly manipulating their initial access to the resource, we operationalized equality as the resulting payoff distribution from the CPR at the end of the game. Thus, children at the age of six are capable of implementing fair distributions in a CPR dilemma to effectively minimize resource competition and thereby increase their collective success in sustaining the resource.

A post-hoc examination of the types of self-distracting conversations that emerged in both conditions revealed a large variety of topics and behaviours. Noteworthy examples are: kissing, playing rock-paper-scissors, singing synchronously, personal stories, riddles and jokes. Across the two conditions, a total of 34 dyads elicited self-distracting (that is, game-irrelevant) utterances. Twelve of these dyads spontaneously began counting synchronously to pass the time, and another three dyads quizzed each other on arithmetic, counting out sums. These counting examples highlight the distracting quality that game-irrelevant utterances took on, and the synchronous nature of the counting exemplifies how self-distraction strategies may have helped to mitigate the difficulty of the delay-of-gratification task. This provides evidence that children are able to socially harness the beneficial effect of self-distraction[47] in an interdependent delay-of-gratification task, using one another as distracting agents to achieve dyadic success. Furthermore, communication among actors in a social dilemma is known to increase the likelihood of cooperation (for example, ref. [14]). Adult studies suggest that the mechanism behind this may be that communication promotes group identity in addition to providing an opportunity to make promises and rules[12]. Additionally, behavioural synchrony has been shown to increase interpersonal affiliation[53] as well as success in joint-action tasks[54]. The self-distraction measure may therefore have served to promote dyadic identity and, in some cases, affiliation and joint action, supporting collective success in addition to the verbal strategy measure.

Various other social factors have been purported to affect behaviour in social dilemmas, such as inter-individual dominance differences[34], trust[55] and social value orientations[56,57]. Further investigation into the role of these interpersonal factors on the development of successful CPR strategies in children would provide a valuable context within which to interpret human CPR behaviour more generally. Results reported by ref. [58] show that increasing group size decreases group members' motivation to divide a shared CPR equally. Future research on this topic should therefore extend CPR paradigms beyond the dyad into the development of group-level dynamics. Last, a cross-cultural exploration of this experimental dilemma would also elucidate how different cultural norms of, for example, reciprocity, prosociality, distributive justice and fairness affect resource distribution in a CPR, and how culture affects the development of children's strategies.

Overall, this study provides evidence that already by the age of six, dyads of children are capable of spontaneously implementing inclusive verbal strategies, equally distributing rewards and socially self-distracting to improve success in sustaining a CPR dilemma. While avoiding the tragedy of the commons is possible, doing so while working independently is much easier than sustaining a resource when a collective dilemma is present. The role of various social and ecological factors that hinder or support success and how sensitivity to these factors develops remain open questions for future experiments.

## Methods

**Participants.** Fifty-four pairs of gender-matched six-year-olds (108 children; mean age: 6 years and 3.3 months; 46 females) were tested. Five additional pairs were excluded: one due to apparatus failure, two for cheating, one for shyness, and one pair were coincidentally already acquainted before testing. Pairs were unfamiliar upon arrival but were familiarized together in a warm-up phase before testing. Children came mainly from middle-class backgrounds in Leipzig, Germany, and were recruited from a database of parents who volunteered to take part in child development studies and came largely from middle-class backgrounds.

**Ethics.** This study was approved by the Max Planck Institute for Evolutionary Anthropology Ethics Committee (members of the committee are M. Tomasello, K. Haberl (head of the child studies laboratory) and research assistant J. Jurkat). The full procedure of the study was covered by the committee's approval. Informed written consent was obtained from all the parents of the children who participated in this study.

**Apparatus.** Dyads were presented with a renewing magic water apparatus for collecting individual egg rewards (see Fig. 1). Blue magic water (hereafter: water) began in a transparent source at the top and flowed via a pump at a steady rate into a transparent cylinder where it was accessible to children through hoses connected to the cylinder bottom. Participants could open and close individual taps on their hoses to control the flow of water into their individual egg reward boxes, which consisted of six transparent cylinders containing one egg each. When the water flowed into the boxes it filled one tube at a time, allowing the eggs to float up to the top one by one. A red cork floated inside the main apparatus cylinder, visually marking the water level and its movement. When one child opened their tap fully to collect water, the outflow was faster than the rate of flow into the cylinder via the pump. When this occurred, the cork dropped with the dropping water level in the cylinder and when it reached a clearly marked red threshold line (20 cm from the bottom of the cylinder; approximately 27% of the lower length of the cylinder) an automatic magnet-based mechanism caused a plug to open at the bottom of the cylinder allowing all the water in the cylinder to drop into a bucket below that was inaccessible to children. Any water continuing to drip from the pump also flowed directly into the bucket. If no collapse occurred, or if children successfully sustained their magic water long enough to access all of the source water before collapsing the resource, the pumping process took 14 min to pump 100% of the water into the cylinder.

**Procedure.** *Introduction.* Children were given a bracelet with either a tree or a sun image and were brought into the testing room. In front of the apparatus were two zones clearly marked with matching tree and sun images, demarcated by white lines on the ground. In each zone there was a seat, a table and an egg reward box. Children were directed to the zone that matched the bracelet they had been given. They were told that the zone represented their individual territory and that they were not allowed to leave it or enter another territory during the game (see Fig. 2).

First, the children experienced an individual familiarization phase with the apparatus. While one child was introduced to the apparatus with a first experimenter (E1), the other child went to play a drawing game in a separate room with a second experimenter (E2). E1 first turned on the pump and showed the child how to open and close the tap, demonstrating that the cork would drop when the tap was open and rise when the tap was closed. E1 explained to the child that opening and closing their tap allowed them to control the quantity and timing of water flow for their own reward collection. The child was then instructed to practise opening and closing the tap on their own three times. All children had at least one—but not more than three, depending on their speed—rounds of opening and closing before the cork reached the threshold, causing water collapse. E1 then explained that collapse occurs when the cork reaches the red threshold line, and that the water was gone. E1 drew the child's attention to the continually dripping water from the pump explaining also that this water was no longer able to flow into their egg reward box. Children were told that they could play however they liked because there were no rules as to when or how long the taps could stay open, but it was important to pay attention to the cork during play. Children were then asked four comprehension questions to reiterate the affordances of the apparatus (see Supplementary Information for details).

*Test.* Dyads participated in only one of the two conditions. Dyads in the collective condition were brought back into the testing room after both had been familiarized with the apparatus. E2 showed them their individual egg collection cylinders, marked with the same tree or sun images to correspond with each participant's bracelet and zone identity, and explained that after the game they could put all of their floating eggs into their egg collection cylinders. E2 also explained individually to each child that for every egg they collected in their egg collection cylinder, they would receive one gummy candy. Both children were instructed to sit at their tables in their zones and were reminded not to leave their zone or enter their partner's during the game. E1 explained that both experimenters had to leave the room after the pump was turned on but that the game would officially begin when E1 gave a double thumbs-up signal. E1 then turned on the pump, gave the thumbs-up signal, and both experimenters left the room.

The procedure in the parallel condition was identical with the exception that instead of one magic water apparatus accessible to both participants simultaneously, each child had access to an individual magic water apparatus in front of them. Both apparatuses in the parallel condition functioned independently so even if one child collapsed their apparatus, their partner could continue to play. In both conditions, there was enough water in one full apparatus to raise six to seven eggs in total (depending on the specific distribution of water). In both conditions, each child always had six available eggs in their egg reward boxes. For dyads in the parallel condition, maximal egg collection was 12 (6 per child with 1 apparatus each), whereas for dyads in the collective condition maximal collection was 6 to 7 eggs (depending on how the water from one apparatus was distributed between partners). The water pumps flowed at the same rate, regardless of condition. See Fig. 2 for the setup.

It should be noted that there was twice the available water per child in the parallel condition compared with the collective condition. While this had no effect on our measure of success, which was proportional to each condition's dyadic affordances, this does mean that where each child had enough water individually in the parallel condition to theoretically collect all their own eggs, pairs shared this same amount of water in the collective condition and had to decide how to allocate the resource between two individuals. The available water and its rate of renewal was held constant per apparatus, renewing at the same steady rate in both conditions. This method was designed after the following example to ensure that the CPR model remained ecologically valid: two equal-sized populations of fish reproduce and regenerate their populations in two separate lakes. These populations would regenerate at the same rate if being fished separately by two independent fishers (as in our parallel condition) or simultaneously by two neighbouring fishers (as in our collective condition). In other words, a resource should be comparable at the outset of a CPR in terms of its quantity and renewal rate regardless of the number of users accessing the resource.

Each dyad played three rounds, with a brief (approximately 3 min) drawing activity in an adjacent room between games so E1 could refill the apparatus(es). Games ended after collapse in the collective condition, after the second of two collapses in the parallel condition, or 30 s after maximal eggs had been collected without collapse. All dyads were surreptitiously filmed from a hidden camera so as to minimize any effects of authority or experimenter expectations on play behaviour.

**Coding and analysis.** E1 coded all videos for egg collection counts per child. A second coder, blind to the study's predictions, re-coded 20% of the videos for egg counts. There was complete agreement between coders for this measure. All trials were transcribed by two coders for verbal dialogue from trial start to collapse or, in the case of no collapse, the end of the trial after the full source water had been pumped into the cylinder and maximum eggs had been collected. Utterances were first categorized as being game-relevant, game-irrelevant or unknown/other. Game-relevant utterances were further categorized as: (1) descriptive, (2) individual strategy, (3) collective strategy, (4) agreement or (5) rejection. All categorizable (that is, game-relevant and game-irrelevant) utterances were used to derive three measures: self-distraction, verbal equality and verbal strategy level. Verbal equality (functionally inequality) was measured as the difference in the number of utterances produced by each child in a dyad per trial over the total number of utterances per trial (denominators for all proportional measures represented as an offset term in statistical models). Self-distraction was defined as the proportion of game-irrelevant utterances out of total dyadic utterances per trial. Verbal strategy level was coded on a 1–4 range for the degree to which children's strategies involved collectively inclusive decisions; trials were assigned the code level of the highest level reached in that trial. Both coders first coded the same 20% subset of trials to compare inter-rater reliability. Agreement was very high between the two verbal coders for self-distraction (Pearson correlation coefficient, $r=0.98$), verbal inequality ($r=0.99$) and verbal strategy level (weighted Cohen's $\kappa=0.973$, $z=5.34$, $P>0.001$). Each coder then proceeded to transcribe and code half the remaining trials.

We designed the coding scheme to reflect a distinction between verbal strategies that pertain to the individual, such as imperatives and descriptive proposals about one's own behaviour, and strategies that pertain to the dyad, such as rules that govern both individuals' behaviour equally and are negotiated collectively. With a coding scheme that differentiates collectively inclusive from individualistic strategies, we predict we can indirectly identify normative mechanisms insofar as we can differentiate when dyads are prescribing behavioural rules that set expectations about the future behaviour of both players, as opposed to when dyads are assigning behavioural expectations unilaterally. Verbal codes and strategy-level definitions can be found in the Supplementary Information. See Table 1 for strategy-level descriptions and examples.

Because the payoff potential differed between the two conditions as a function of water availability at the start of the trial coming from one or two apparatuses, we looked at the proportion of available eggs collected by the dyad as our dependent measure. There was enough water available at the start of the trial for dyads with minimum success to collect one egg in the collective condition and four eggs in the parallel condition, even when collapse occurred as quickly as possible. For this reason, success was measured at the trial level by the number of eggs collected by dyads in the collective condition minus one and the number of eggs collected by dyads in the parallel condition minus four. Likewise, when dyads were maximally successful in sustaining the CPR—when collapse occurred after all the water had dripped from the source into the game cylinder—dyads could collect a maximum of 7 eggs in the collective condition and 12 eggs in the parallel condition. After correcting these maximum-success amounts per condition for minimum success, we used the following formula to calculate a success measure, comparable across conditions, of the proportion of available eggs gained by the dyad:

$$\text{Corrected success} = \frac{(\text{number of dyadic eggs} - \text{minimum success per condition (1/4)}}{(\text{maximum available eggs per condition (7/12)} - \text{minimum success per condition (1/4))}}$$

All statistical models presented were run in R (R Core[59]) using the 'glmer' function for generalized linear mixed models[60] of the R package 'lme4'[61].

First, we analysed dyadic success ($n=162$ trials; 87 in the collective condition, 75 in the parallel condition) between the two conditions and across trials 1–3 with a Poisson model including condition and trial number as test predictors, as well as their interaction. We controlled for the effects of sex and included an offset term in the model for the total number of available eggs depending on condition (collective: 7; parallel: 12).

Next, we tested for differences between the two conditions on the behavioural measures of verbal equality and proportional self-distraction using three Poisson models. We also compared the likelihood of reaching terminal verbal strategy (level 4) between the two conditions using a binomial model ($n=141$ trials). These four models analysed the effect of a single test predictor—condition—on the behavioural measures as responses, controlling for sex and trial number, and including offset terms for proportions where appropriate. Verbal equality comprised two models. One model included all trials for which audio data were available (21 trials removed from the full dataset for all further analyses due to microphone failure) minus all trials with a dyadic total of zero utterances before collapse ($n=122$). A second model was run on the same dataset after removing all trials with complete inequality ($n=102$). The proportional self-distraction model sample size was 134 trials after removing all trials containing zero total dyadic utterances and zero game-irrelevant utterances.

To investigate how success was achieved in the collective condition, we ran two further Poisson models with success as the dependent measure, determined in the collective condition by the number of eggs collected by the dyad, controlling for sex and trial number. In the first model ($n=72$ collective condition trials) we included the test predictors of equality of eggs collected by partners within dyads (1 – egg difference between partners / total dyadic eggs collected) and verbal strategy level (1–4). The total number of eggs available in the collective condition was six to seven (depending on water distribution between partners), and the egg equality measure ranged from 0–1. Egg equality reflected differences of 0 eggs for equal dyads (with a distribution of 1:1, 2:2 or 3:3) to egg differences of 5–6 for maximally unequal dyads (with egg distributions of 6:1 or 6:0, respectively). The second model was run with a slightly reduced dataset ($n=51$ collective condition trials excluding trials with zero categorizable utterances) that included the test predictors of verbal equality and proportion self-distraction. See Supplementary Information for all model descriptions, results and statistical details.

**Reporting Summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

**Code availability.** R scripts that support the statistical methods presented for plotted results in this study are available in the Supplementary Information. Specific sourced functions in the R code are available from the corresponding author upon request.

**Data availability.** Verbal data that support the findings of this study are available from the corresponding author upon request. All original egg collection data can be found in the Supplementary Information.

Received: 8 September 2017; Accepted: 23 February 2018;
Published online: 16 April 2018

## References

1. Ostrom, E., Gardner, R. & Walker, J. *Rules, Games, and Common-Pool Resources* (Univ. Michigan Press, Ann Arbor, MI, 1994).
2. Hardin, G. The tragedy of the commons. *Science* **162**, 1243–1248 (1968).
3. Tomasello, M. *Why We Cooperate* (MIT Press, Cambridge, MA, 2009).
4. Acheson, J. M. *The Lobster Gangs of Maine* (Upne, Lebanon, NH, 1988).
5. Anderson, L. R., Mellor, J. M. & Milyo, J. Inequality and public good provision: an experimental analysis. *J. Socio-Econ.* **37**, 1010–1028 (2008).
6. Tavoni, A., Dannenberg, A., Kallis, G. & Löschel, A. Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proc. Natl Acad. Sci. USA* **108**, 11825–11829 (2011).
7. Ostrom, E. The value-added of laboratory experiments for the study of institutions and common-pool resources. *J. Econ. Behav. Organ.* **61**, 149–163 (2006).
8. Hackett, S., Schlager, E. & Walker, J. The role of communication in resolving commons dilemmas: experimental evidence with heterogeneous appropriators. *J. Environ. Econ. Manag.* **27**, 99–126 (1994).
9. Schmitt, P., Swope, K. & Walker, J. Collective action with incomplete commitment: experimental evidence. *S. Econom. J.* **66**, 829–854 (2000).
10. Janssen, M. A. Introducing ecological dynamics into common-pool resource experiments. *Ecol. Soc.* **15**, 8 (2010).
11. Biel, A. & Thøgersen, J. Activation of social norms in social dilemmas: a review of the evidence and reflections on the implications for environmental behaviour. *J. Econ. Psychol.* **28**, 93–112 (2007).
12. Orbell, J. M., Van de Kragt, A. J. & Dawes, R. M. Explaining discussion-induced cooperation. *J. Pers. Soc. Psychol.* **54**, 811–819 (1988).
13. Richerson, P. & Henrich, J. Tribal social instincts and the cultural evolution of institutions to solve collective action problems. *Cliodynamics* **3**, 38–80 (2009).
14. Dawes, R. M., McTavish, J. & Shaklee, H. Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *J. Pers. Soc. Psychol.* **35**, 1–11 (1977).
15. Dawes, R. M., Orbell, J. M., Simmons, R. T. & Van De Kragt, A. J. Organizing groups for collective action. *Am. Political Sci. Rev.* **80**, 1171–1185 (1986).
16. Croson, R. T. Theories of commitment, altruism and reciprocity: evidence from linear public goods games. *Econ. Inq.* **45**, 199–216 (2007).
17. Goldstein, N. J., Cialdini, R. B. & Griskevicius, V. A room with a viewpoint: using social norms to motivate environmental conservation in hotels. *J. Consum. Res.* **35**, 472–482 (2008).
18. Kazemi, A. & Eek, D. in *New Issues and Paradigms in Research on Social Dilemmas* (eds Biel A., Eek, D., Gärling T. & Gustafsson, M.) 72–92 (Springer, Boston, MA, 2008).
19. Colding, J. & Folke, C. Social taboos: "invisible" systems of local resource management and biological conservation. *Ecol. Appl.* **11**, 584–600 (2001).
20. Cardenas, J.-C. & Carpenter, J. Behavioural development economics: lessons from field labs in the developing world. *J. Dev. Stud.* **44**, 311–338 (2008).
21. Kerr, N. L. in *Social Dilemmas: Social Psychological Perspectives* (ed. Schroeder, D.) 31–47 (Pergamon Press, New York, NY, 1995).
22. Agrawal, A. Common property institutions and sustainable governance of resources. *World Dev.* **29**, 1649–1672 (2001).
23. Gummerum, M., Hanoch, Y. & Keller, M. When child development meets economic game theory: an interdisciplinary approach to investigating social development. *Hum. Dev.* **51**, 235–261 (2008).
24. Wellman, H. M., Cross, D. & Watson, J. Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* **72**, 655–684 (2001).
25. Sally, D. & Hill, E. The development of interpersonal strategy: autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* **27**, 73–97 (2006).
26. Takagishi, H., Kameshima, S., Schug, J., Koizumi, M. & Yamagishi, T. Theory of mind enhances preference for fairness. *J. Exp. Child Psychol.* **105**, 130–137 (2010).

27. Feinfield, K. A., Lee, P. P., Flavell, E. R., Green, F. L. & Flavell, J. H. Young children's understanding of intention. *Cogn. Dev.* **14**, 463–486 (1999).
28. Call, J. & Tomasello, M. Distinguishing intentional from accidental actions in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *J. Comp. Psychol.* **112**, 192–206 (1998).
29. Sodian, B., Taylor, C., Harris, O. L. & Perner, J. Early deception and the child's theory of mind: false trails and genuine markers. *Child Dev.* **62**, 468–483 (1991).
30. Grueneisen, S., Wyman, E. & Tomasello, M. Conforming to coordinate: children use majority information for peer coordination. *Br. J. Dev. Psychol.* **33**, 136–147 (2015).
31. Grueneisen, S., Wyman, E. & Tomasello, M. "I know you don't know I know…" Children use second-order false-belief reasoning for peer coordination. *Child Dev.* **86**, 287–293 (2015).
32. Sloane, S., Baillargeon, R. & Premack, D. Do infants have a sense of fairness? *Psychol. Sci.* **23**, 196–204 (2012).
33. Blake, P. R., McAuliffe, K. & Warneken, F. The developmental origins of fairness: the knowledge–behavior gap. *Trends Cogn. Sci.* **18**, 559–561 (2014).
34. Grueneisen, S. & Tomasello, M. Children coordinate in a recurrent social dilemma by taking turns and along dominance asymmetries. *Dev. Psychol.* **53**, 265–273 (2017).
35. Vogelsang, M., Jensen, K., Kirschner, S., Tennie, C. & Tomasello, M. Preschoolers are sensitive to free riding in a public goods game. *Front. Psychol.* **5**, 729 (2014).
36. Harbaugh, W. T. & Krause, K. Children's altruism in public good and dictator experiments. *Econ. Inq.* **38**, 95–109 (2000).
37. Camerer, C. *Behavioral Game Theory* (New Age International, New Delhi, 2010).
38. House, B., Henrich, J., Sarnecka, B. & Silk, J. B. The development of contingent reciprocity in children. *Evol. Hum. Behav.* **34**, 86–93 (2013).
39. Sanchez-Amaro, A., Duguid, S., Call, J. & Tomasello, M. Strategic decision-making by chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*) and children in a snowdrift-game task. Preprint at https://doi.org/10.7287/peerj.preprints.1805v1 (2016).
40. Kortenkamp, K. V. & Moore, C. F. Time, uncertainty, and individual differences in decisions to cooperate in resource dilemmas. *Pers. Soc. Psychol. Bull.* **32**, 603–615 (2006).
41. Mischel, W., Shoda, Y. & Rodriguez, M. L. Delay of gratification in children. *Science* **244**, 933–938 (1989).
42. Mischel, W. & Ebbesen, E. B. Attention in delay of gratification. *J. Pers. Soc. Psychol.* **16**, 329–337 (1970).
43. Yates, G. C., Yates, S. M. & Beasley, C. J. Young children's knowledge of strategies in delay of gratification. *Merrill Palmer Q.* **33**, 159–169 (1987).
44. McCabe, L. A. & Brooks-Gunn, J. With a little help from my friends?: Self-regulation in groups of young children. *Infant Ment. Health J.* **28**, 584–605 (2007).
45. Nisan, M. Delay of gratification in children: personal versus group choices. *Child Dev.* **47**, 195–200 (1976).
46. Dawes, R. M., Ven De Kragt, A. J. & Orbell, J. M. Not me or thee but we: the importance of group identity in eliciting cooperation in dilemma situations: experimental manipulations. *Acta Psychol.* **68**, 83–97 (1988).
47. Mischel, W., Ebbesen, E. B. & Raskoff Zeiss, A. Cognitive and attentional mechanisms in delay of gratification. *J. Pers. Soc. Psychol.* **21**, 204–218 (1972).
48. Vaughn, B. E., Kopp, C. B., Krakow, J. B., Johnson, K. & Schwartz, S. S. Process analyses of the behavior of very young children in delay tasks. *Dev. Psychol.* **22**, 752–759 (1986).
49. Rodriguez, M. L., Mischel, W. & Shoda, Y. Cognitive person variables in the delay of gratification of older children at risk. *J. Pers. Soc. Psychol.* **57**, 358–367 (1989).
50. Messick, D. M. et al. Individual adaptations and structural change as solutions to social dilemmas. *J. Pers. Soc. Psychol.* **44**, 294–309 (1983).
51. Köymen, B., Schmidt, M. F., Rost, L., Lieven, E. & Tomasello, M. Teaching versus enforcing game rules in preschoolers' peer interactions. *J. Exp. Child Psychol.* **135**, 93–101 (2015).
52. Bardhan, P. & Dayton-Johnson, J. in *The Drama of the Commons* 87–112 (National Academies Press, Washington DC, 2002).
53. Hove, M. J. & Risen, J. L. It's all in the timing: interpersonal synchrony increases affiliation. *Soc. Cogn.* **27**, 949–960 (2009).
54. Valdesolo, P., Ouyang, J. & DeSteno, D. The rhythm of joint action: synchrony promotes cooperative ability. *J. Exp. Soc. Psychol.* **46**, 693–695 (2010).
55. Mulder, L. B., Van Dijk, E., De Cremer, D. & Wilke, H. A. Undermining trust and cooperation: the paradox of sanctioning systems in social dilemmas. *J. Exp. Soc. Psychol.* **42**, 147–162 (2006).
56. Balliet, D., Parks, C. & Joireman, J. Social value orientation and cooperation in social dilemmas: a meta-analysis. *Group Process. Inter. Relat.* **12**, 533–547 (2009).
57. Parks, C. D. The predictive ability of social values in resource dilemmas and public goods games. *Pers. Soc. Psychol. Bull.* **20**, 431–438 (1994).
58. Allison, S. T., McQueen, L. R. & Schaerfl, L. M. Social decision making processes and the equal partitionment of shared resources. *J. Exp. Soc. Psychol.* **28**, 23–42 (1992).
59. R Development Core Team R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2016).
60. Baayen, R. H. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R* (Cambridge Univ. Press, Cambridge, 2008).
61. Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: Linear mixed-effects models using Eigen and S4. R Package Version 1 (2014).

## Acknowledgements

This work was supported by the Max Planck Society. We thank the research support staff at the Department of Developmental and Comparative Psychology child studies laboratory for help with data collection, including N. Bobovnikov. We thank R. Mundry for statistical oversight and C. Stephens for statistical instruction. We thank J. Lang and R. Hagedorn for efforts in transcribing and coding the verbal data. We thank C. Piot for apparatus drawings and A. C. Schneider for helpful comments on the manuscript. Special thanks go to R. Pieszek and S. Schütte for help designing and building the apparatus. Finally, we thank all the children and families who participated in the study. The funders had no role in study design, data collection, data analysis, decision to publish or preparation of the manuscript.

## Author contributions

R.K. and E.H. contributed to the study design. R.K. collected the data. R.K. and E.H. analysed the data. R.K. and E.H. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41562-018-0327-2.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Correspondence and requests for materials** should be addressed to R.K.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Rebecca Koomen

☐ Initial submission ☐ Revised version ☒ Final submission

# Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

## Experimental design

1. Sample size

| | |
|---|---|
| Describe how sample size was determined. | The sample size of 54 dyads (29 in collective condition; 25 in parallel condition) was determined without a power analysis based on previous research involving commons dilemmas, as well as previous work involving delay of gratification in children. |

2. Data exclusions

| | |
|---|---|
| Describe any data exclusions. | Five dyads of children were excluded from the dataset: 1 for apparatus failure, 2 for cheating, 1 for shyness, and 1 for being previously acquainted prior to coming to the institute for testing. Exclusion criteria were established prior to data collection. |

3. Replication

| | |
|---|---|
| Describe whether the experimental findings were reliably reproduced. | No attempt was made by the authors to replicate the original findings presented in the manuscript. |

4. Randomization

| | |
|---|---|
| Describe how samples/organisms/participants were allocated into experimental groups. | Children were randomly assigned to dyads based on age, gender, and availability. All dyads were randomly assigned to a test condition prior to their arrival at the institute. |

5. Blinding

| | |
|---|---|
| Describe whether the investigators were blinded to group allocation during data collection and/or analysis. | The experimenter was not blind to condition during data collection as the setup involved a different number of apparatuses per condition. Reliability coding was done by behavioural coders blind to all predictions and results of the study. |

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.) |
| ☐ | ☒ | A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | A statement indicating how many times each experiment was replicated |
| ☐ | ☒ | The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section) |
| ☐ | ☒ | A description of any assumptions or corrections, such as an adjustment for multiple comparisons |
| ☐ | ☒ | The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted |
| ☐ | ☒ | A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range) |
| ☒ | ☐ | Clearly defined error bars |

*See the web collection on statistics for biologists for further resources and guidance.*

## Software

Policy information about availability of computer code

7. Software

| | |
|---|---|
| Describe the software used to analyze the data in this study. | All statistical models presented were run in R (R Core Team) using the glmer function for Generalized Linear Mixed Models (GLMM) of the R package lme4. |

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

## Materials and reagents

Policy information about availability of materials

8. Materials availability

| | |
|---|---|
| Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company. | No unique materials were used in the study. |

9. Antibodies

| | |
|---|---|
| Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species). | No antibodies were used. |

10. Eukaryotic cell lines

| | |
|---|---|
| a. State the source of each eukaryotic cell line used. | No eukaryotic cell lines were used. |
| b. Describe the method of cell line authentication used. | No eukaryotic cell lines were used. |
| c. Report whether the cell lines were tested for mycoplasma contamination. | No eukaryotic cell lines were used. |
| d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use. | No commonly misidentified cell lines were used. |

## Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

| | |
|---|---|
| Provide details on animals and/or animal-derived materials used in the study. | No animals were used in this study. |

Policy information about studies involving human research participants

12. Description of human research participants

| Describe the covariate-relevant population characteristics of the human research participants. | Fifty-four gender-matched pairs of children ranging from 6 years to 6 years and 6 months were tested in this study. All pairs were unfamiliar before testing and came from similar socio-economic backgrounds in a mid-sized German city. |
| --- | --- |