Reciprocity and the tragedies of maintaining and providing the commons

Simon Gächter1,2,3*, Felix Kölle4 and Simone Quercia5

Social cooperation often requires collectively beneficial but individually costly restraint to maintain a public good2–5, or it needs costly generosity to create one6–8. Status quo effects4 predict that maintaining a public good is easier than providing a new one. Here, we show experimentally and with simulations that even under identical incentives, low levels of cooperation (the ‘tragedy of the commons’2) are systematically more likely in maintenance than provision. Across three series of experiments, we find that strong and weak positive reciprocity, known to be fundamental tendencies underpinning human cooperation29,30, are substantially diminished under maintenance compared with provision. As we show in a fourth experiment, the opposite holds for negative reciprocity (‘punishment’). Our findings suggest that incentives to avoid the ‘tragedy of the commons’ need to contend with dilemma-specific reciprocity.

Humans are an exceptionally cooperative species able to collaborate for the creation of common benefit31–33. Collective actions such as voting, participating in political movements, the provision of the welfare state, charity, volunteering and teamwork are examples of public goods that come into existence by the generosity of many people who put the greater good before self-interest2–4. Cooperation is, however, not always about providing collectively valuable resources, but often about maintaining existing ones6–8. Limiting carbon dioxide emissions, sustaining natural resources or maintaining common pastures and biodiversity are important examples of cooperation problems that require restraint in exploiting existing socially beneficial public goods.

In this paper, we show experimentally and with simulations that cooperation for maintaining an initially existing public good is substantially and systematically weaker than cooperation for creating a new public good even if they are otherwise identical social dilemmas. This is unexpected, given that many people are biased towards the status quo and defaults9, which should ease cooperation when the public good already exists compared with when it needs to be provided.

We show that the reason for lower cooperation in the maintenance dilemma is that reciprocity, a fundamental force behind the evolution of cooperation and human sociality34–36, is substantially diminished in maintaining compared with providing a public good. Simulations show that despite some variability, lower cooperation in maintenance than provision is a systematic effect to be expected with a likelihood of 70%. The simulation results also provide an explanation for the mixed findings in some related literature14–23.

In our experiments, we focus sharply on the behavioural differences between initially existing and inexistant public goods (Fig. 1) and abstract from technological complexities, loss aversion, time discounting and institutional details relevant in real-world social dilemmas2,23–29. In maintenance, a group of four people possesses a common pool of 80 tokens and each member can withdraw up to 20 tokens. Upholding the status quo by withdrawing nothing earns each group member 32 money units (MU); if all people withdrew maximally, everyone earns 20 MU. In provision, the common pool is initially empty and 80 tokens are distributed equally among group members who decide simultaneously how many tokens (up to 20) to contribute to the pool. In the status quo, all earn 20 MU, and if all people contribute maximally, each member earns 32 MU.

Using the setup described in Fig. 1, we ran three series of experiments with 704 participants who interacted anonymously in three generic settings of social interaction (see Methods). All experiments involved a between-subjects comparison of cooperation in maintenance and provision. We also elicited beliefs about group members’ contributions to (or withdrawals from) the public good. Participants needed to successfully complete a comprehension test before the experiment started.

In the first experiment, called ‘one-shot’, participants (n = 288) made a single decision only. This experiment is a basic measure of people’s cooperativeness in the absence of strategic incentives to cooperate. In the second experiment, called ‘strangers’, participants (n = 256) played the games from Fig. 1 for 27 iterations while the group composition was randomly changed in each round. This experiment was a sequence of one-shot interactions that permitted learning about cooperativeness in the population30–33. The third experiment (n = 160), called ‘partners’, kept the group composition constant across the 27 iterations, which created strategic incentives for cooperation32,33.

The effective size of the public good (after withdrawals or contributions) was smaller in maintenance than provision in all the experiments (Fig. 2 and Supplementary Table 1). In one-shot, the public good in maintenance was, on average, 27% smaller than in provision (Fig. 2, panel 1; 23.8 versus 32.6, respectively; two-sided t-test, t = −2.51, P = 0.014). Low levels of the public good (less than 10% of the optimal size of 80) were more likely in maintenance than provision (23% versus 0%, respectively; chi-squared test, χ²(1) = 9.51, P = 0.002).

In strangers, the public good started out 23% lower in maintenance than provision (22.7 versus 29.5, respectively; two-sided t-test, t = −1.92, P = 0.059) and decayed on average to about 5% of the socially efficient level in both problems (Fig. 2, panel 2). Thus, the tragedy of the commons was almost maximal in both maintenance and provision.
In partners, the public good started 33% smaller in maintenance than provision (27.7 versus 41.3, respectively; two-sided t-test, t = −2.96, P = 0.005) and dropped over time (Fig. 2, panel 3). On average, the public good was 37.3% smaller in maintenance than provision (10.6 versus 16.9, respectively; linear mixed effects model, P = 0.035).

Comparing partners and strangers revealed the extent to which strategic incentives help the provision of the public good. We found that in maintenance the average size of the public good was only 3.6 units higher in partners than strangers (10.6 versus 7.0, respectively; linear mixed effects model, P = 0.346; Supplementary Table 2), while in provision the public good was on average twice as large in partners than strangers (16.9 versus 8.2, respectively; linear mixed effects model, P = 0.004). Thus, strategic incentives to increase cooperation were substantially weaker in maintenance than provision.

Taken together, these results show that high levels of the public good are harder to achieve in maintenance than provision in one-shot and in the first period of partners and strangers. This is surprising given that in maintenance the public good enjoys a head start because it is already provided at the outset. Furthermore, while in strangers the size of the public good converges to similar long-run equilibrium levels, in partners the initial differences are persistent and lead to different long-run outcomes between provision and maintenance. The aim of our further analysis was to understand the differences in cooperation outcomes by investigating the maintenance or provision dilemma. The elicited attitudes are separating empirically whether initial resource allocation affects reciprocation.

Studying reciprocation as a preference requires looking beyond cooperation outcomes and measuring attitudes to cooperation separately from outcomes. The reason why this is important is that people who differ in their ex ante attitudes can make the same cooperation decision ex post. To see why, consider that a conditional cooperators’s ex ante attitude is to cooperate only if they believe their group members do so too. However, there may also be ‘free riders’, who never want to contribute to the public good irrespective of their beliefs regarding how much others contribute. A conditional cooperators believe that others do not contribute and a person with a free-rider attitude both contribute nothing: their ex post behaviour is observationally equivalent despite different ex ante attitudes. Thus, if cooperation is a function of attitudes and beliefs, the challenge is to separate them empirically. Our approach, which we call the ‘ABC of cooperation’, achieves this separation. This also allows us to compare strong reciprocation as measured by the ABC approach with reciprocation estimated from observed behaviour.

The ABC approach measures individual attitudes (a), beliefs (b) and effective contributions (c) separately and explains cooperation as a(b) → c. It is inspired by ref. 30 and implemented as follows. All three experiments start with an incentive-compatible elicitation of attitudes without feedback in a one-shot version of either the maintenance or provision dilemma. The elicited attitudes are our main measure of strong reciprocation. Eliciting attitudes involves specifying a vector a, of contributions or withdrawals as a function of all possible average contributions or withdrawals of other group members. We classify participants as conditional cooperators (that is, strong reciprocators) if the entries in the vector a, is increasing in others’ contributions or withdrawals, or as a free rider if a participant’s a, consists of only zero contributions or maximal withdrawals. We refer to the remaining participants as ‘others’. After attitude elicitation, the three experiments proceed as described above. In all experiments, we elicit incentivized beliefs (b) about other group members’ average withdrawal or contribution and we observe effective contributions (c) to the public good (see Methods).

In the repeated direct interactions of strangers and partners we measure conditional cooperation in linear mixed effects models by
regressing individual contributions or withdrawals on the average contribution or withdrawals of other group members in the previous period (Supplementary Information). The relation between these two variables, the coefficient $\beta_i$, is our measure of conditional cooperation. We call $\beta_i$, ‘estimated reciprocity’.

In strangers, $\beta_i$ is an estimate of strong reciprocity because there are no strategic incentives to pretend being a reciprocator. As $\beta_i$ is estimated from behaviour only, it is a proxy for strong reciprocity. However, we expect that participants with attitudes that classify them as conditional cooperators have $\beta_i > 0$, whereas people with a free rider attitude display $\beta_i \approx 0$.

In partners, conditional cooperators also have $\beta_i > 0$, which may be larger in strangers due to the added incentives for weak reciprocity. Free riders may therefore also display $\beta_i > 0$. Furthermore, we use the attitudes $\alpha_i$ and $\beta_i$ to study the link between strong and weak reciprocity.

Elicited attitudes were significantly different in maintenance and provision ($\chi^2(16) = 31.03, P < 0.001$; Fig. 3a). In maintenance, participants were significantly less likely to be conditional cooperators than in provision (42% versus 64%, respectively; $\chi^2(1) = 31.03, P < 0.001$), significantly more likely to be free riders (28% versus 17%, respectively; $\chi^2(1) = 10.46, P = 0.001$) and significantly more likely to display an unclassified attitude (‘others’); 30% versus 19%, respectively; $\chi^2(1) = 11.08, P = 0.001$). Thus, in maintenance 58% of participants did not reciprocate their group member’s effective contribution, which is almost the mirror image of the 64% in provision who did reciprocate.

Estimated reciprocity $\beta_i$ in the repeated games was also significantly lower in maintenance than provision in both strangers and partners (Fig. 3b, panel 1; multilevel mixed effects models, $P < 0.001$; Supplementary Table 3). The added strategic incentives for weak reciprocity significantly increased the estimated reciprocity in both maintenance and provision (Fig. 3b, panel 1; multilevel mixed effects models, $P < 0.001$; Supplementary Table 4).

Estimated reciprocity was also consistent with attitude types elicited before the repeated games (Supplementary Tables 5 and 6). Participants classified as conditional cooperators showed high degrees of estimated reciprocity in strangers and partners, which was significantly higher than that of free riders in both maintenance and provision (Fig. 3b, panels 2 and 3; multilevel mixed effects models, $P < 0.001$). Conditional cooperators also displayed significantly higher $\beta_i$ in partners than strangers (multilevel mixed effects models, $P < 0.001$). As predicted, free riders in strangers displayed low estimated reciprocity but showed increased $\beta_i$ in partners compared with strangers. Participants classified as ‘others’ displayed a substantial $\beta_i$ but did not react to strategic incentives (Fig. 3b, panel 4; multilevel mixed effects models, $P > 0.166$).

Our next step was to investigate whether the differences in reciprocity across maintenance and provision could explain the observed differences in cooperation outcomes (Fig. 2). We did this by applying our ABC framework, which uses attitudes and beliefs to explain effective contributions. We calculated predicted effective contributions $a_i(b_i) \rightarrow e_i$ and compared them with actual effective contributions $c_i$ from one-shot as well as the effective first-period contributions in the repeated experiments (Methods). Predicted and actual effective contributions were highly significantly positively correlated in one-shot as well as in all repeated games (all Spearman’s $p > 0.59$; $P < 0.001$).

We also calculated individual-level deviations from the predicted effective contribution, $e_i \rightarrow c_i$. In one-shot, this measure lay within ±2 tokens in 63% and 62% of the cases in maintenance and provision, respectively, with no differences between treatments ($\chi^2(1) = 0.01, P = 0.903$). We obtained similar results for first-period effective contributions in strangers (66% and 63%, respectively; $\chi^2(1) = 0.43, P = 0.514$) and partners (74% and 64%, respectively; $\chi^2(1) = 1.86, P = 0.172$). Finally, effective contributions differed significantly between attitude types: free riders contributed significantly less than conditional cooperators and ‘others’ in all conditions (Supplementary Fig. 1).

The fact that the ABC approach predicted equally well in maintenance and provision allowed us to use the elicited attitudes and beliefs as a ‘population pool’ from which to sample at random to run ‘simulated experiments’ (Methods and Supplementary Information). The advantage of simulations was that we were not restricted to a specific laboratory sample we happened to draw at a given instance (with hitherto unobservable attitudes and beliefs); we could perform a large number of identical experiments cost effectively and therefore elicit a distribution of likely cooperation.

**Fig. 3 | Reciprocity in maintenance and provision.** a. Strong reciprocity as measured by cooperation attitudes; type classification as in ref. 30, $n_s = 348$; $n_p = 356$. Chi-squared tests, **$P < 0.01$. The results are robust to alternative classification methods (Supplementary Information). b. Estimated reciprocity in repeated interactions (±1 standard error of the mean; strangers (S); $n = 256$; partners (P); $n = 160$) by treatment and attitude category (conditional cooperators, free riders or others). Positive reciprocity was estimated as the coefficient of lagged average contributions of the other group members ($\bar{C}_{i,t-1}$) from multilevel mixed effects linear regressions (Supplementary Information and Supplementary Table 3). An alternative estimation approach using finite mixture models15 confirmed these results (Supplementary Table 14).
The results of 1,000 simulated experiments (Fig. 4) showed that effective cooperation levels in maintenance were lower than in provision in 70% of all simulated experiments. This result shows that our findings that cooperation in maintenance was lower than in provision is systematic.

Given that our results revealed important asymmetries in positive reciprocity between maintenance and provision, it was interesting to study whether initial resource allocation also affected negative reciprocity, which in our setting took the form of punishment. Furthermore, punishment is an expression of moral disapproval and social norms that are important in many real-world public goods. If the differences in positive reciprocity in maintenance and provision also translate into negative reciprocity, we should observe less punishment in maintenance than provision and, therefore, also a reduced effectiveness of punishment to stabilize cooperation in maintenance compared with provision.

We studied punishment in a fourth experiment (‘partners with punishment’; \(n = 172\)), which was identical to partners except for an added punishment stage in each period after group members had made their withdrawal or contribution decisions. In the punishment stage, each group member could assign up to five punishment points to each other member, where each punishment point cost one MU and reduced the earnings of the punished group member by three MU (see Methods).

The attitudes elicited before the experiment replicated the results from Fig. 3a (Supplementary Fig. 2). Contrary to expectations, negative reciprocity, estimated as assigned punishment in reaction to negative deviations of others from one’s own effective contribution, was substantially and significantly higher in maintenance than provision. This effect was present both overall and for each attitude type (Fig. 5a, panels 1–4) and was not driven by different frequencies of punishers (Supplementary Fig. 3). There were no treatment differences for positive deviations (Fig. 5a, panel 1 and Supplementary Table 7). Interestingly, in contrast to the estimated positive reciprocity, the estimated negative reciprocity did not differ between conditional cooperators and free riders (Fig. 5a, panels 2–4 and Supplementary Table 8).

As expected, punishment increased the public goods to substantially higher levels compared with partners (Fig. 5b; linear mixed effects models; maintenance: 43.1 versus 10.6, respectively, \(P < 0.001\); provision: 44.1 versus 16.9, respectively, \(P < 0.001\); Supplementary Tables 9 and 10). Remarkably, the sizes of public goods were very similar in maintenance and provision (linear mixed effects models; maintenance: 43.1, provision: 44.1, \(P = 0.904\)). Besides stronger negative reciprocity, a further reason for this result was that reactions to received punishment (in terms of a change in effective contributions) were also stronger in maintenance than provision (Supplementary Table 11).

One way to reconcile the results on positive and negative reciprocity in partners and partners with punishment, respectively, is to argue that also in partners people engage in punishment by reducing their contributions in the current period as a reaction to previous negative deviations of others from one’s own effective contributions. If such ‘implicit’ punishment were stronger in maintenance than provision, it could explain why the decay in effective contributions was stronger in maintenance than provision. However, this conjecture is not borne out by the data.

We found that participants in partners significantly increased their contributions in round 1 in response to positive deviations of others from their own contributions in round 1-1; the reverse held for negative deviations. However, we found both of these reactions to be significantly more pronounced in provision than maintenance (linear mixed effects models; both \(P < 0.018\); Supplementary Table 12). This confirms, once again, stronger conditional cooperation in provision compared with maintenance in partners. It also suggests another interpretation of the results of partners with punishment: because voluntary conditional cooperation was weaker in maintenance than provision, stronger extrinsic incentives were needed—here, in the form of punishment—to stabilize cooperation in maintenance at similar levels to provision.

Our analysis revealed that the important principles of human cooperation of strong and weak reciprocity were substantially diminished when cooperation required restraint in exploiting a public good as opposed to when cooperation called for generosity to provide a public good. Our findings are consistent with the observation that failing to contribute to a public good is judged more morally blameworthy than exploiting an existing public good.

Our results can also be explained by a model of revealed altruism, according to which initial resource allocation affects perceptions of generosity of actions and hence subsequent reciprocity. Because in provision cooperation is the result of an act of commission (contributing), while in maintenance cooperation is achieved by omission (not withdrawing), cooperation in provision is perceived as more generous than in maintenance and thus provision triggers stronger positive reciprocity than maintenance. In contrast, our results suggest that negative reciprocity, as expressed by people’s costly punishment, does not follow this logic because punishment is more severe in maintenance than provision, and likely to compensate for weaker voluntary cooperation in maintenance.

Our findings from the experiments without punishment and the simulations also help to explain the mixed evidence from previous related literature, which, with a few exceptions, only compared cooperation outcomes; that is, the effective size of the public good after contribution or withdrawal decisions. Some of these studies found higher cooperation in so-called ‘give-some’ versus ‘take-some’ games, some found the reverse and some found no significant differences. The simulations based on our ABC approach can explain these mixed results (Fig. 4), but they also show that, on average, cooperation in maintenance is generally expected to be lower than provision. The finding that maintenance and provision are systematically different also suggests that future research should...
choose the game (maintenance or provision) that comes closest to the social dilemma of interest.

Our results also have potential policy relevance\(^4,5\). Recent policy proposals to foster cooperation build on the power of reciprocity in combination with economic incentives\(^6-8\). Policy makers who reckon with reciprocity should therefore consider that the extent of reciprocity that can be evoked is dilemma specific. Moreover, a problem of incentives is that they might ‘crowd out’ strong reciprocity because incentives typically strengthen self-regarding motives to cooperate\(^9\). Our finding of higher reciprocity in provision than maintenance suggests that crowding out may be more problematic in provision problems than maintenance problems, because in maintenance more people display non-reciprocal attitudes. Future research will need to address these issues, including how reciprocity and incentives interact in non-linear settings with thresholds, resource rivalry, discounting and hybrid social dilemmas where provision and exploitation can take place at the same time.

Methods

**Isomorphism of maintenance and provision under monetary incentives.** In maintenance, each group of four members was initially endowed with 80 tokens placed in a ‘group project’; individual members had no endowment. Material incentives are described by equation (1):

\[
\pi_i = w_i + 0.4 \left(80 - \sum_{j=1}^{4} w_j \right)
\]  

(1)

where \(0 \leq w_i \leq 20\) indicates the withdrawal of individual ‘i’ from the project.

In provision, the ‘group project’ was initially empty and each group member had an endowment of 20 tokens instead. The material incentives for each individual ‘i’ are described by equation (2):

\[
\pi_i = 20 - c_i + 0.4 \sum_{j=1}^{4} c_j
\]  

(2)

where \(0 \leq c_i \leq 20\) denotes the contribution of individual ‘i’ to the project.

Hence, under rationality and money maximization, maintenance and provision are isomorphic social dilemmas. Using \(c_i = 20 - w_i\) for \(j = 1, \ldots, 4\) and substituting into equation (2) we obtain equation (1). Analogously, using \(w_j = 20 - c_i\) for \(j = 1, \ldots, 4\) and substituting into equation (1) yields equation (2).

**Experimental design details.** The experiments were approved by the Research Ethics Committee at the School of Economics, University of Nottingham. We conducted four series of experiments using the two decision situations described above and in Fig. 1. Each experiment was composed of three parts that allowed elicitation of the three components of the ABC framework: an individual’s attitude (\(a_i\)) towards cooperation (\(a_i\)’s ‘type’), \(a_i\)’s beliefs (\(b_i\)) about others’ contributions and \(a_i\)’s contribution decision (\(c_i\)). Participants knew that the experiment consisted of several parts but only received information about the relevant part on progression of the experiment. To avoid spillover effects between different parts, information about decisions and payoffs was given only at the very end of the experiment. Experimental instructions are in the Supplementary Information.

In part 1, participants were introduced to either the maintenance or provision problem. Before continuing, participants answered a set of computerized control questions.

In part 2, we elicited cooperation attitudes \(a_i\), using a variant of the strategy method\(^10\), which allowed elicitation of an individual’s willingness to cooperate as a function of the other group members’ cooperation decisions. Participants were asked to make an unconditional and conditional cooperation decision. In the unconditional decision, they were simply asked how much they wanted to withdraw from (or contribute to) the common pool. In the conditional contribution, participants had to fill a withdrawal (or contribution) table in which they had to state their withdrawal (or contribution) decision for each possible (rounded) average withdrawal (or contribution) of the other three group members. This gave us the vector \(a_{ij}\), our measure of strong reciprocity. To achieve incentive compatibility, in each group a random mechanism selected three members for which the unconditional decision was payoff relevant and one member for whom the conditional decision for the (rounded) average unconditional withdrawal (or contribution) of the three other group members was payoff relevant.

Part 3 comprised a direct-response interaction that differed in its exact design protocol across the four experiments as described in the main text (one-shot, strangers, partners and partners with punishment). This elicited component \(c_i\) of the ABC framework.

In all repeated experiments (strangers, partners and partners with punishment), participants were matched in groups of four and interacted for 27 consecutive rounds under payment rules of equation (1) or equation (2). Participants were not told how many rounds the experiment would last for\(^11\). This avoided endgame effects and also seemed realistic for many common resource problems, which do not have a known endpoint. In strangers, participants were re-matched randomly in matching

---

**Fig. 5** | **Partners with punishment.** a. Estimated negative reciprocity (±1 standard error of the mean) by treatment and attitude category (conditional cooperators, free riders or others). We estimated negative reciprocity in multilevel mixed effects linear regression as the number of punishment points assigned to effective contributions that deviated negatively from one’s own contribution (Supplementary Table 7 and Supplementary Information). b. Effective levels of the public goods (±1 standard error of the mean) over the 27 rounds of interactions (\(n_p = 84; n_o = 88\)).
groups of 16 participants each after every round, while in partners and partners with punishment the group composition remained constant across all 27 rounds. At the end of each round, participants received aggregate feedback on choices and outcomes. In all rounds of the direct-response interactions, we also elicited beliefs about average effective contributions of the other three group members. Participants were paid for the accuracy of their beliefs. They earned three points if their belief was exactly correct, two points when their belief deviated by one point from the true average effective contribution and one point when their belief deviated by two points. If their estimation was off by more than two points, they received no additional money. This elicited component b of our framework.

Data collection and subject-pool socio-demographics. A total of n = 876 students participated in the experiments (maintenance: nM = 432; provision: nP = 444). Participants were recruited with the help of the Online Recruitment System for Economic Experiments (ref. 2) from the volunteer student subject pool at the University of Nottingham. Participants gave informed consent. The average age was 20.1 years (standard deviation: 2.25 years) and 57% were females. Nationalities were 59% British, 22% Asian, 12% from other European countries and the rest from other countries. The subjects studied included 20% economics or business, 18% other social sciences, 20% humanities, 14% sciences, 12% engineering, 12% medical science and 4% law. We conducted all experiments in the Centre for Decision Research and Experimental Economics laboratory at the University of Nottingham using z-Tree (ref. 3). The experiments lasted between 70 and 210 min depending on the experimental condition. Participants earned on average £20.60.

Predicting effective contributions. In one-shot, strangers and partners, the ABC approach allowed us to predict contributions using elicited cooperation attitudes a, and beliefs b; b = b (h) → c. By matching beliefs with the corresponding decision in the contribution (or withdrawal) table, we predicted a contribution (or withdrawal) decision c for each subject and compared c with the actual contribution c that we observed in the direct-response experiment24-25.

Classification of attitudes. We analysed cooperation in the strategy-method experiment treating each participant’s effective contribution schedule (the vector a) as an independent observation. We classified cooperation attitudes into three main behavioural types: (1) ‘condition’ cooperators, whose effective contribution schedule exhibited a (weakly) monotonically increasing pattern or whose Spearman’s correlation coefficient between their schedule and the others’ average contribution was positive and significant at the 1% level; (2) ‘free riders’, who never contributed anything (always withdrew everything) irrespective of how much the others contributed (or withdrew); or (3) ‘others’, when neither (1) nor (2) applied. Attitudes were a proxy for cooperation preferences because they reflected a willingness to pay for cooperation as a function of other group members’ cooperation.

Simulations. For each simulated experiment, we randomly sampled (with replacement) from the participant pool of maintenance experiments attitudes and beliefs (n = 60, median sample size in related studies also using linear public goods26) and calculated simulated effective contributions a (h) b, c. We did the same for n = 60 provision attitudes and beliefs. This resembled an experiment in which a researcher invites 60 participants per treatment and then observes their effective contribution. As a participation pool, we used all n = 876 attitudes from our four experiments (maintenance: nM = 432; provision: nP = 444) and n = 544 beliefs (maintenance: nM = 268; provision: nP = 276) from one-shot as well as the first period of strangers. See Supplementary Information for details.

Statistical analysis. In the one-shot direct interaction, we treated (b and c) as independent observations. In the repeated interactions, we treated beliefs and effective contributions at the matching group level as an independent observation. Matching groups comprised 16 participants in strangers and four each participants in partners and partners with punishment. For the repeated experiments, all estimations were performed using linear mixed models with random intercepts at the matching group and the individual level (see Statistical analysis in Supplementary Information for details of the model specifications).

Code availability. We used Stata version 14.2 (http://www.stata.com/) for the data analyses. The codes are stored in a Dryad data package with the title ‘Reciprocity in maintaining and providing public goods’ and can be found at http://dx.doi.org/10.5061/dryad.8d9r2.

Data availability. The data that support the findings of this study are stored in a Dryad data package with the title ‘Reciprocity in maintaining and providing public goods’ and can be found at http://dx.doi.org/10.5061/dryad.8d9r2.

Received: 8 March 2017; Accepted: 28 July 2017; Published online: 28 August 2017

References
1. Van Lange, P. A. M., Balliet, D., Parks, C. D. & Van Vugt, M. Social Dilemmas. The Psychology of Human Cooperation (Oxford Univ. Press, Oxford, 2014).
2. Hardin, G. The tragedy of the commons. Science 162, 1243–1248 (1968).
3. Ostrom, E. Governing the Commons. The Evolution of Institutions for Collective Action (Cambridge Univ. Press, Cambridge, 1990).
4. Collier, P. The Plundered Planet (Penguin Books, London, 2011).
5. Gintis, H., Bowles, S., Boyd, R. & Fehr, E. Moral Sentiments and Material Interests. The Foundations of Cooperation in Economic Life (MIT Press, Cambridge, 2005).
6. Samuelson, W. & Zeckhauser, R. Status quo bias in decision making. J. Risk Uncertain. 1, 7–59 (1988).
7. Trivers, R. L. The evolution of reciprocal altruism. Q. Rev. Biol. 46, 35–57 (1971).
8. Axelrod, R. & Hamilton, W. The evolution of cooperation. Science 211, 1390–1396 (1981).
9. Bowles, S. & Gintis, H. A Cooperative Species: Human Reciprocity and its Evolution (Princeton Univ. Press, Princeton, 2011).
10. Rand, D. G. & Nowak, M. A. Human cooperation. Trends Cogn. Sci. 17, 413–425 (2013).
11. Nowak, M. A. & Highfield, R. Supercooperators: Altruism, Evolution, and Why We Need Each Other to Succeed (Free Press, New York, 2011).
12. Turchin, P. Utopia or Notopia: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth (Beresta Books, Chaplin, CT, 2016).
13. Tomasetti, M. A. Natural History of Human Morality (Harvard Univ. Press, Cambridge, MA, 2016).
14. Andreoni, J. Warm glow versus cold prickle—the effects of positive and negative framing on cooperation in experiments. Q. J. Econ. 110, 1–21 (1995).
15. Park, E.-S. Warm-glow versus cold-prickle: a further experimental study of framing effects on free-riding. J. Econ. Behav. Organ. 43, 405–421 (2000).
16. Khadiv, M. & Lange, A. Doing good or doing harm: experimental evidence on giving and taking in public good games. Exp. Econ. 18, 432–441 (2015).
17. Cox, C. A. Decomposing the effects of negative framing in linear public goods games. Econ. Lett. 126, 63–65 (2015).
18. Sell, J. & Son, Y. Comparing public goods and common pool resources: three experiments. Soc. Psychol. Q. 60, 118–137 (1997).
19. Van Dijk, E. & Wilke, H. Is it mine or is it ours? Framing property rights and decision making in social dilemmas. Organ. Behav. Hum. Decis. Process. 71, 195–209 (1997).
20. Cubitt, R., Drouvelis, M. & Gächter, S. Framing and free riding: emotional responses and punishment in social dilemma games. Exp. Econ. 14, 254–272 (2011).
21. Dufwenberg, M., Gächter, S. & Hennig-Schmidt, H. The framing of games and the psychology of play. Games Econ. Behav. 73, 459–478 (2011).
22. Messer, K. D., Zarghamee, H., Kaiser, H. M. & Schulze, W. D. New hope for the voluntary contributions mechanism: the effects of context. J. Public Econ. 91, 1783–1799 (2007).
23. Cox, J. C., Ostrom, E., Sadiraj, V. & Walker, J. M. Provision versus appropriation in symmetric and asymmetric social dilemmas. South. Econ. J. 79, 496–512 (2013).
24. Aposteguia, J. & Maier-Rigaud, P. F. The role of rivalry. Public goods versus common pool resources. J. Confl. Resolut. 50, 646–663 (2006).
25. Levin, S. A. Public goods in relation to competition, cooperation, and spite. Proc. Natl Acad. Sci. USA 111, 10838–10845 (2014).
26. Brewer, M. B. & Kramer, R. M. Choice behavior in social dilemmas: effects of social identity, group size, and decision framing. J. Pers. Soc. Psychol. 50, 543–549 (1986).
27. De Dreu, C. K. & McCusker, C. Gain–loss frames and cooperation in two-person social dilemmas: a transformational analysis. J. Pers. Soc. Psychol. 72, 1093–1106 (1997).
28. Van Dijk, E., Wilke, H., Wilke, M. & Meteman, L. What information do we use in social dilemmas? Environmental uncertainty and the employment of coordination rules. J. Exp. Soc. Psychol. 35, 109–135 (1999).
29. McCusker, C. & Carnevale, P. J. Framing in resource dilemmas: loss aversion and the moderating effects of sanctions. Organ. Behav. Hum. Dec. Process. 61, 190–201 (1995).
30. Fischbacher, U. & Gächter, S. Social preferences, beliefs, and the dynamics of free riding in public good experiments. Am. Econ. Rev. 100, 541–556 (2010).
31. Janssen, M. A., Holaha, R., Lee, A. & Ostrom, E. Lab experiments for the study of social-ecological systems. Science 328, 613–617 (2010).
32. Kessler, C. & van Winden, F. Conditional cooperation and voluntary contributions to public goods. Scand. J. Econ. 102, 23–39 (2000).
33. Andreoni, J. Why free ride—strategies and learning in public-goods experiments. J. Public Econ. 37, 291–308 (1988).
34. Rustagi, D., Engel, S. & Kosfeld, M. Conditional cooperation and costly monitoring explain success in forest commons management. Science 330, 961–965 (2010).
35. Fosgaard, T. R., Hansen, L. G. & Wengström, E. Understanding the nature of cooperation variability. *J. Public Econ.* **120**, 134–143 (2014).

36. Gintis, H. Strong reciprocity and human sociality. *J. Theor. Biol.* **206**, 169–179 (2000).

37. Fehr, E. & Fischbacher, U. Social norms and human cooperation. *Trends Cogn. Sci.* **8**, 185–190 (2004).

38. Nyborg, K. et al. Social norms as solutions. *Science* **354**, 42–43 (2016).

39. Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).

40. Sigmond, K. Punish or perish? Retaliation and collaboration among humans. *Trends Ecol. Evol.* **22**, 593–600 (2007).

41. Cubitt, R., Drouvelis, M., Gächter, S. & Kabalin, R. Moral judgments in social dilemmas: how bad is free riding? *J. Public Econ.* **95**, 253–264 (2011).

42. Cox, J. C., Friedman, D. & Sadiraj, V. Revealed altruism. *Econometrica* **76**, 31–69 (2008).

43. Cox, J. C., Servátka, M. & Vadovič, R. Status quo effects in fairness games: reciprocal responses to acts of commission versus acts of omission. *Exp. Econ.* **20**, 1–18 (2017).

44. Fleishman, J. A. The effects of decision framing and others' behavior on cooperation in a social dilemma. *J. Confl. Resolut.* **32**, 162–180 (1988).

45. Frackenpohl, G., Hillenbrand, A. & Kube, S. Leadership effectiveness and institutional frames. *Exp. Econ.* **19**, 842–863 (2016).

46. Fehr-Duda, H. & Fehr, E. Game human nature. *Nature* **530**, 413–415 (2016).

47. MacKay, D. J. C., Cramton, P., Ockenfels, A. & Stoft, S. Price carbon—I will if you will. *Nature* **326**, 315–316 (2015).

48. Rand, D. G., Yolli, E. & Hoffman, M. Harnessing reciprocity to promote cooperation and the provisioning of public goods. *Policy Insights Behav. Brain Sci.* **1**, 263–269 (2014).

49. Bowles, S. Policies designed for self-interested citizens may undermine "the moral sentiments": evidence from economic experiments. *Science* **320**, 1605–1609 (2008).

50. Fischbacher, U., Gächter, S. & Fehr, E. Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* **71**, 397–404 (2001).

51. Rand, D. G. et al. Positive interactions promote public cooperation. *Science* **325**, 1272–1275 (2009).

52. Greiner, B. Subject pool recruitment procedures: organizing experiments with ORSEE. *J. Econ. Sci. Assoc.* **1**, 114–125 (2015).

53. Fischbacher, U. Z-tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* **10**, 171–178 (2007).

54. Fischbacher, U., Gächter, S. & Quercia, S. The behavioral validity of the strategy method in public good experiments. *J. Econ. Psychol.* **33**, 897–913 (2012).

55. Bardsley, N. & Moffatt, P. G. The experimenters of public goods: inferring motivations from contributions. *Theory Decis.* **62**, 161–193 (2007).

**Acknowledgements**

We thank A. Alonso Arechar, B. Beranek, T. Cason, J. Cox, F. Fallucchi, U. Fischbacher, M. García-Vega, R. Hernán-González, D. Houser, L. Molleman, D. van Dolder, T. Weber, O. Weisel and various seminar audiences for helpful comments. B. Beranek provided valuable research assistance. This work was supported by the European Research Council grant ERC-AdG 295707 COOPERATION and the Economic and Social Research Council Network for Integrated Behavioural Sciences (ES/K002201/1). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

**Author contributions**

S.G., F.K. and S.Q. developed the research ideas and designed the study. F.K. and S.Q. conducted the experiments and analysed data. S.G., F.K. and S.Q. wrote the manuscript.

**Competing interests**

The authors declare no competing interests.

**Additional information**

Supplementary information is available for this paper at doi:10.1038/s41562-017-0191-5.

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

Correspondence and requests for materials should be addressed to S.G.

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