# Supporting Information

Rockenbach and Milinski 10.1073/pnas.1108996108

## SI Methods

One hundred twenty undergraduate students from the University of Erfurt voluntarily participated in 24 experimental sessions [eight sessions of each of the three treatments, i.e., a simple public goods game (SPG), a public goods game with punishment possibility (PG&PUN), and a public goods game followed by an indirect reciprocity game (PG&IR)], with five participants in each session. These sessions constituted the independent observations for the nonparametric statistical analysis. Special care was exerted to recruit students from many different disciplines to increase the likelihood that the subjects had never met before. Each participant was allowed to take part in only one session.

The experiment consisted of 30 rounds of play, organized in two equal-length blocks of 15 rounds. Before the first round of block 1 began, one of the subjects was chosen randomly to be the observer. The remaining four subjects were the players. Roles remained fixed for the entire first block; it was known that the first block's observer would become a player for the second block. At the beginning of the second block, the first block's observer chose three the players to be her fellow players for the second block. The participant not chosen as a player became the observer of the second block. First-block observers had the choice of having the second-block observer determined by a random draw or by actively selecting the second-block players. Active selection incurred a cost of 0.5 points to the first-block observer.

Each of the 30 rounds of play had an identical structure. In SPG the round consisted only of stage S1, in which a linear public goods game was played. In PG&PUN and in PG&IR each round consisted of stage S1, in which the public goods game was played, and stage S2 with a peer punishment possibility (in PG&PUN) or with an indirect reciprocity game (in PG&IR).

**Public Goods Game in Stage S1 of SPG, PG&PUN, and PG&IR.** The public goods game constitutes a prototypical social dilemma in which each player is endowed with 10 points and may contribute 0, 5, or 10 points to a public good. The points contributed to the public good are doubled and create a benefit for the entire group. This group benefit is distributed equally among the group members (i.e., each member profits by 0.5 points from each 1 point contributed; 0.5 is the marginal per capita return). The points not contributed to the public good are transferred to the participant's private account. Because the cost of contributing (e.g., 1 point) is higher than the individual return from that contribution (e.g., 0.5 points), it always is in the material self-interest of any participant to free-ride on the contributions of the others and to keep all points for the private account. If all participants follow their material self-interest, no one contributes to the public good, and each participant achieves a payoff of 10 points. Because the group's return of each point invested is greater than 1, it is in the collective interest that all group members contribute their entire endowment to the group project. These diametrically opposed individual and collective interests constitute the social dilemma in public good provision. At the end of stage 1, after the players have simultaneously made their contribution decisions, they are informed about the contributions of the other players.

**Peer Punishment Possibility in Stage S2 of PG&PUN.** At the beginning of stage S2 in PG&PUN, each player receives 3 additional points. These points may be used to punish the other players. Each player must decide whether to allocate one punishment point to each of the other players. A punishment point reduces the punisher's account by 1 point and the punished player's account by 3 points. The points not used for punishment are transferred to the player's private account.

**Indirect Reciprocity Game in Stage S2 of PG&IR.** At the beginning of stage S2 in PG&IR, each player receives 3 additional points. These points may be given to another player, the "receiver." Points given to another player are tripled, so that the receiver can receive 9 points. If the sender opts not to send, the sender keeps the 3 points, and nothing reaches the receiver. The sender is informed about the contribution the receiver made in stage S1 of the actual round. Each player acts both as a sender and a receiver in each round. To exclude direct reciprocity, the sender and the receiver are matched to assure that the receiver has not previously been the sender's sender and will not be in the future. This provision was made absolutely clear to the participants.

**Observation Decision.** Players could choose to hide actions, and the observer could choose to observe players' actions. Before each stage the observer and the four players simultaneously made their decisions to observe or allow observations, respectively. Observers could observe, at most, two players' decisions. For each of the four players the observer had to decide whether she would (*i*) not observe a player's decision, (*ii*) observe a player's decision openly, or (*iii*) observe the player's decision without the player's knowledge (concealed condition). When not observing, the observer was not informed about the player's decision. When observing openly or concealed, the observer was informed about the player's decision, if the player allowed that decision to be observed (see below). When the observer observed openly, the player was informed that the observer would be informed about the player's decision. When the observer observed concealed, the player was not informed. Concealed watching incurred a cost of 0.5 points to the observer; open watching and not observing incurred no costs. Each player had to decide whether to allow his decisions to be observed (open window) or to prohibit observation (closed window). When the window was closed, there was no way for the observer to be informed about the player's decision. Closing the window incurred a cost of 0.5 points to the player; an open window incurred no costs. The observation decisions of the observer and the players had to be made before each stage of the round (i.e., before the contribution decision in S1 and before the decision to punish or give, respectively, in S2).

At the end of the round the observer and each player was informed about the personal payoff in the round. The observer received 100 points for the entire experiment and an endowment of 2 points for each round and had to bear the costs of observation and selection costs incurred. A player's payoff was the sum of the payoffs in S1 and S2.

At the beginning of the experiment participants received written and oral instructions. At the end of the experiment participants privately received their experimental earnings in cash at a conversion rate of 0.05 € per 1 point. All experimental decisions were made on a computer screen using the experimental software z-Tree (1).

**Play of a Money-Maximizing Player.** In a world of players interested solely in maximizing their monetary gains, each player follows his dominant strategy of the one-shot game: he keeps his entire endowment for his private account, does not invest in punishing other players in PG&PUN, does not give in PG&IR, and does not invest any money to close the window. This strategy is im-

mediately clear for the second block, where there is no observer choice at its end. However, a money-maximizing observer who is aware that all players free-ride in the second block has no incentive to bear any costs to select any particular players for the second block. Consequently, she opts for a random selection of the second block's observer. Anticipating this option, the players do not have any incentives to deviate from the dominant strategy of the one-shot game. Hence, if it is common knowledge that all participants are engaging in money-maximizing behavior, we do not observe any contributions, any punishment, or any giving. Furthermore, the players do not invest in hiding their actions, and the observer does not invest in seeking information while concealed.

**Companion Study with Exogenous Observation Choice.** In a companion study we investigated how observation decisions that are not made endogenously by the players but are enforced exogenously by the experimenter influence behavior. The companion study was done with fresh participants from the same subject pool.

The design of the experiment is identical to the one reported in this article, apart from two modifications. (*i*) A random draw decides whether a player's window is open (probability of 2 in 3) or closed (probability of 1 in 3). (*ii*) When observers decide to observe a player, a random draw determines whether this observation is open (probability of 1 in 2) or concealed (probability of 1 in 2). The probability distributions are common knowledge, and the random draws for all players and all observers at all stages of the game (contribution and punishment) are independent. Note that the players' decisions about contribution and punishment are not determined randomly but, as in the present study, are made by the individual participants. Similarly, the observer's decision of whom to observe, whether to determine the next observer actively or randomly, and—in the former case—whom to select are not determined randomly but are made by the individual observers, as in the present study.

The decisions concerning contribution and punishment that observers witnessed were very similar to those of the present study. When all window decisions are made endogenously by the subjects, the average contribution in PG&PUN is 8.50, and the average number of players punished is 0.32. With exogenous window decisions the average contribution in PG&PUN is 8.48, and the average number of players punished is 0.13.

Therefore, we could rerun regression 1 (Table 1, *Column B*) for the extended data set from 32 independent observations, incorporating the 24 observations on the endogenous window choice and the eight data sets on exogenous window determination in PG&PUN (Table S1). The regression strongly confirms the results obtained with the smaller data set. The difference between the lowest and the second-lowest contributor is the significant determinant on the observer's decision to choose the next observer actively ($P = 0.003$). When this difference is large, the probability that the observer will actively choose the next observer increases significantly. The difference between the lowest and the second-lowest punisher, which is weakly significant (at 0.079) when relying only on the endogenous-choice data, is not significant in the larger data set ($P = 0.731$). Fig. S1 graphically shows the selection decision of the observer in the new data set and is analogous to Fig. 2 in the main text.

1. Fischbacher U (2007) z-Tree. Zurich toolbox for ready-made economic experiments. *Exp Econ* 10:171–178.

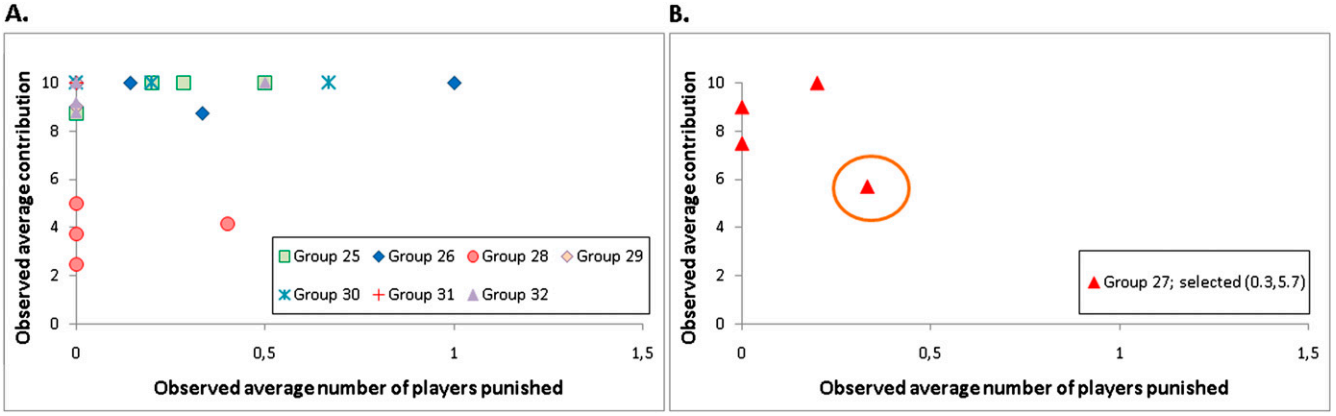


**Fig. S1.** Determination of the next observer in the companion study with a random determination of the players' window status (open or closed). Average of number of coplayers punished and average contributions of the individual players the observer observed in the first block. Each symbol represents an individual player and designates the group. (*A*) Groups in which the observer opted for a random choice of next block's observer. (*B*) Groups in which the observer actively chose next block's observer. The circled player is the one selected as next observer.

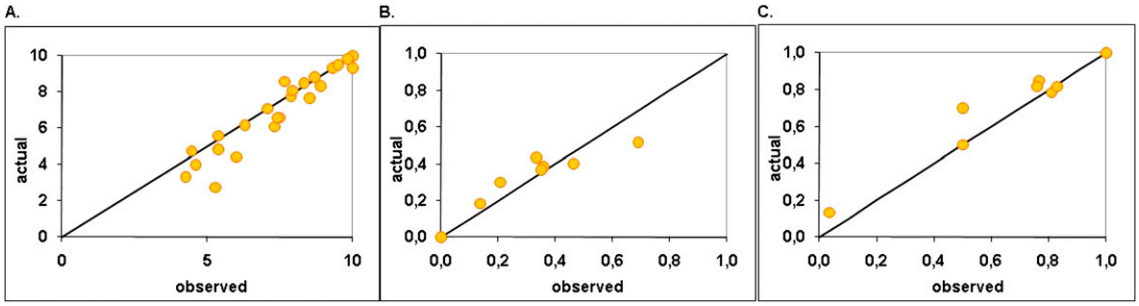


**Fig. S2.** Observed and actual behavior. (*A*) Average of observed contributions versus average of all contributions per group. (*B*) Average of observed number of punished players versus average of the number of all punished players. (*C*) Average of observed decisions to give versus average of all decisions to give. For statistics see main text.

**Table S1. Probit regression (present and companion study combined) for the probability that the observer actively chooses the observer of the second block**

| | Coefficient | *z* | *P* |
|---|---|---|---|
| Average contribution of all players in block 1 over the 15 rounds of block 1 | −0.0829699 | −0.47 | 0.640 |
| Difference between the second-lowest and the lowest contributor in block 1 | 0.8352608 | 2.93 | 0.003 |
| Difference between the highest and the second-highest punisher in block 1 | −0.4389293 | −0.36 | 0.718 |
| Difference between the lowest and the second-lowest punisher in block 1 | −1.699731 | −0.34 | 0.731 |
| Treatment dummy PG&PUN | −1.076936 | −1.39 | 0.165 |
| Constant | −0.2204929 | −0.15 | 0.885 |

Phase 1 only; averages over the values the observer actually has seen (either openly or concealed); 32 independent observations; robust SEs, clustered for independent observations.

**Table S2. Probit regression for probability of having the window open for the decisions to contribute/not contribute, punish/not punish, and give/not give**

| | Decision to contribute/not contribute (A)* | | | Decision to punish/not punish (B)† | | | Decision to give/not give (C)‡ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coefficient | *z* | *P* | Coefficient | *z* | *P* | Coefficient | *z* | *P* |
| Number of past periods with an open window in this decision node | 0.4320706 | 10.96 | 0.000 | 0.4730158 | 2.74 | 0.006 | 0.4581372 | 13.70 | 0.000 |
| Contribution | 0.1802546 | 6.52 | 0.000 | −0.0808108 | −1.62 | 0.105 | 0.0651397 | 1.70 | 0.090 |
| Sum of past contributions | −0.0192695 | −3.34 | 0.001 | −0.0000835 | −0.01 | 0.995 | 0.0014663 | 0.23 | 0.821 |
| Allocated punishment points | | | | −0.6959632 | −3.87 | 0.000 | | | |
| Sum of past allocated punishment points | | | | 0.0937431 | 1.59 | 0.113 | | | |
| Decision to give | | | | | | | 1.19343 | 2.21 | 0.027 |
| Sum of past decisions to give | | | | | | | −0.1577813 | −2.36 | 0.018 |
| Period | −0.2658565 | −9.82 | 0.000 | −0.4486862 | −1.88 | 0.060 | −0.3224914 | −7.47 | 0.000 |
| Constant | 0.9636747 | 5.09 | 0.000 | 3.211178 | 3.78 | 0.000 | 1.031011 | 3.15 | 0.002 |

Phase 1 only; robust SEs; clustered for independent observations.
*Over 24 independent observations.
†Over eight independent observations.
‡Over eight independent observations.

**Table S3. Probit regression for probability that the concealed observer observes the player at the decisions to contribute/not contribute, punish/not punish, and give/not give**

| | Decision to contribute/not contribute (A)* | | | Decision to punish/not punish (B)† | | | Decision to give/not give (C)‡ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coefficient | *z* | *P* | Coefficient | *z* | *P* | Coefficient | *z* | *P* |
| Number of times the observer has observed the player in this decision node so far | −0.0652034 | −1.12 | 0.262 | | | | −0.0068536 | −0.13 | 0.896 |
| Number of times the observer wanted to observe the player, but the window was closed in this decision node so far | 0.1042304 | 1.01 | 0.315 | | | | | | |
| Average contributions the observer has observed so far | 0.0738779 | 2.15 | 0.032 | | | | 0.092278 | 0.94 | 0.348 |
| Average allocated punishment points the observer has observed so far | | | | | | | | | |
| Average decisions to give the observer has observed so far | | | | | | | 0.0797544 | 0.15 | 0.878 |
| Period | −0.0125647 | −0.39 | 0.699 | | | | −0.0692426 | −2.83 | 0.005 |
| Constant | −2.011884 | −6.89 | 0.000 | | | | −2.059941 | −1.90 | 0.058 |

Phase 1 only; (robust SEs; clustered for independent observations.
*Over 24 independent observations.
†Over eight independent observations. There were too few data to run regression.
‡Over eight independent observations.

**Table S4. Probit regression for probability that the observer observes the player openly at the decisions to contribute/not contribute, punish/not punish, and give/not give**

| | Decision to contribute/ not contribute (A)* | | | Decision to punish/not punish (B)† | | | Decision to give/not give (C)‡ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coefficient | *z* | *P* | Coefficient | *z* | *P* | Coefficient | *z* | *P* |
| Number of times the observer has observed the player in this decision node so far | 0.0375462 | 0.65 | 0.519 | −0.0838511 | −0.84 | 0.403 | 0.1671187 | 3.85 | 0.000 |
| Number of times the observer wanted to observe the player, but the window was closed in this decision node so far | 0.0946006 | 1.22 | 0.224 | −0.250468 | −1.16 | 0.247 | 0.4130242 | 1.38 | 0.169 |
| Average contributions the observer has observed so far | −0.0338983 | −2.68 | 0.007 | 0.0537368 | 1.81 | 0.070 | −0.0369251 | −0.92 | 0.360 |
| Average allocated punishment points the observer has observed so far | | | | 0.0512202 | 0.29 | 0.774 | | | |
| Average decisions to give the observer has observed so far | | | | | | | −0.1615924 | −0.91 | 0.362 |
| Period | −0.0105726 | −0.42 | 0.677 | 0.0294581 | 0.60 | 0.551 | −0.0655107 | −2.23 | 0.026 |
| Constant | −0.0051711 | −0.04 | 0.966 | −0.3974666 | −2.13 | 0.033 | 0.1007889 | 0.25 | 0.801 |

Phase 1 only; (robust SEs; clustered for independent observations.
*Over 24 independent observations.
†Over eight independent observations.
‡Over eight independent observations.