Supporting Information

Riedl et al. 10.1073/pnas.1203179109

SI Materials and Methods

Subjects. Details on the subjects, namely their ages, sex, and rearing histories are provided (Table S7). The numbers of groups tested based on kinship with the victim and dominance to the thief are shown in Table S8. Table S9 provides details on the 60 test groups.

Test groups. Sixty test groups were created based on dominance relationships between actor and thief and kinship relationships between actor and victim.

SI Results and Discussion

Following are details on the generalized linear mixed model (GLMM) results that were presented in the article (Tables S1–S5).

Whole Sample: Simplifying the Model. For the whole sample of 60 test groups, there was an initial full model consisting of four types of variables. The response variable; three random effects factors: actor, thief, and victim; test variables that defined the experimental design: condition, dominance, and kinship; and control variables: actor's age, sex, rearing history (nursery vs. mother-reared), session order, and trial order per session.

Full model: response ~ condition + dominance + kinship + age + sex + rearing + session order + trial order within session + actor + thief + victim.

The coefficients were estimated by using maximum likelihood, and we used a binominal error structure and logit link function. A likelihood ratio test (using the ANOVA function in R) was used to measure the change of the fit between the full model and a reduced model (not comprising the factor of interest). A significant finding indicated that the factor in question contributed to the variance of the response variable "collapsing the trapdoor."

To guard against false-positives (1), we tested whether the whole set of test variables defining the experimental design significantly explained variance in the response variable by testing the full model against the null model (which did not comprise any of the test factors but included the control predictors). This comparison revealed significance ($\chi^2_7 = 26.927$, $P < 0.001$), which supports the interpretation that the tested factors influenced the response of the actors.

We investigated the potential interaction effects of condition, dominance, and kinship using only the third-party (3P) unfair and 3P theft conditions because these were the only conditions in which both thieves and victims were present. We analyzed the interactions by including them in the original full model. Neither the interactions between condition*dominance nor dominance*kinship were significant (Table S2). Note that the interaction term condition*kinship could not be computed because of low response rates in some conditions in interaction with kinship. A comparison between the full model and the null model showed a trend: $\chi^2_5 = 10.139$, $P = 0.07139$.

Instead of using dominance and kinship (and their interaction) as predictor variables, we also tested the effect of these variables by using the single predictor variable "test group," which consisted of the four categories shown in Table S8 combined into a single factor. A comparison between the full model and the null model (only consisting of control variables) revealed significance: $\chi^2_8 = 26.452$, $P = 0.00087$. The factor group category did not significantly explain variance on the response variable ($\chi^2_3 = 6.4155$, $P = 0.093$).

Analysis of Test Conditions. From Table S1, it can be seen that only the random-effects factor "actor," the control factor "session order," and the experimental test variables "condition" and "dominance" contribute to the variance of the response variable "collapsing the trapdoor." The single effects on the response of each significant control and experimental test factor in the final model are shown in Table S3. Because the test factor "condition" was significant ($\chi^2_5 = 19.508$, $P = 0.002$), post hoc pairwise comparisons between different conditions were performed. We ran the final model twice to use both the second- and the third-party theft condition (2P theft, 3P theft) as reference conditions to show differences between the theft conditions and every other condition.

Final model: response ~ condition + dominance + session order + actor

Dominant Actors. Both, the experimental test variable "condition" and the control variable "session order" were significant predictors (condition: $\chi^2_5 = 17.252$, $P = 0.004$; session order: $\chi^2_1 = 7.8767$, $P = 0.005$). All pairwise comparisons between conditions are displayed in Table S4. We ran the final model twice to use both the second- and the third-party theft condition (2P theft, 3P theft) as reference conditions to display differences between the theft conditions and every other condition.

Model: response ~ condition + dominance + session order + actor

Subordinate Actors. Both the experimental test variable "condition" and the control variable "session order" were significant predictors (condition: $\chi^2_5 = 16.223$, $P = 0.006$; session order: $\chi^2_1 = 3.9659$, $P = 0.046$). All pairwise comparisons between conditions are displayed in Table S5. We ran the final model twice to use both, the second- and the third-party theft condition (2P theft, 3P theft) as reference conditions to display differences between the theft conditions and every other condition.

Model: response ~ condition + dominance + session order + actor

Nonparametric Analyses. Because mixed models are still under development and their estimates a matter of debate (2), we also analyzed the data with Friedman's and Wilcoxon's exact-signed rank tests. All analyses are two-tailed. The results of these analyses support those presented in the article.

Responses of subjects in all groups did not differ across conditions (Friedman's χ^2_5 test = 9.700, $P = 0.084$). Given that there was a trend, we did pairwise comparisons between the conditions. There was no difference between 2P theft and 2P loss [Wilcoxon T^+ test = 23.00, $n = 13$ (five ties), $P = 0.570$], 2P theft and 3P theft [Wilcoxon T^+ test = 41.5, $n = 13$ (two ties), $P = 0.473$], 3P theft and 3P loss [Wilcoxon T^+ test = 10.00, $n = 13$ (six ties), $P = 0.563$], and 3P theft and 3P no victim [Wilcoxon T^+ test = 1.00, $n = 13$ (eight ties), $P = 0.125$]. There was a trend between 3P theft and 3P unfair [Wilcoxon T^+ test = 2.00, $n = 13$ (six ties), $P = 0.063$], but there was no a priori reason to expect this difference which might arise because of multiple comparisons.

Responses of dominant subjects differed across conditions (Friedman's χ^2_5 test = 13.059, $P = 0.017$). Importantly, there was a strong trend for dominants to collapse the platform more often in 2P theft than 2P loss [Wilcoxon T^+ test = 15.00, $n = 12$ (seven ties), $P = 0.062$], consistent with previous findings (3). There was a weak tendency for dominants to collapse the trapdoor more often in 2P theft than 3P theft [Wilcoxon T^+ test = 30.00, $n = 12$ (four ties), $P = 0.102$]. In contrast, dominant individuals did not engage in third-party punishment by collapsing the plat-

form more often in response to 3P theft [3P theft vs. 3P loss: Wilcoxon T^+ test = 6.00, n = 12 (seven ties), P = 0.813; 3P theft vs. 3P unfair: Wilcoxon T^+ test = 3.50, n = 12 (six ties), P = 0.188; 3P theft vs. 3P no victim: Wilcoxon T^+ test = 3.00, n = 12 (seven ties), P = 0.313]. As with the GLMM data, nonparametric tests also showed that subordinates did not collapse the platform in response to the different conditions [Friedman's χ^2_5 test = 7.090, P = 0.217; 2P theft vs. 2P loss: Wilcoxon T^+ test = 3.50, n = 11 (six ties), P = 0.375; 2P theft vs. 3P theft: Wilcoxon T^+ test = 4.00, n = 11 (seven ties), P = 1.000; 3P theft vs. 3P unfair: Wilcoxon T^+ test = 0, n = 11 (10 ties), P = 1.000; 3P theft vs. 3P loss: Wilcoxon T^+ test = 2.00, n = 11 (nine ties), P = 1.000; 3P theft vs. 3P no victim: Wilcoxon T^+ test = 1.00, n = 11 (eight ties), P = 0.500].

Individual Differences. Because a few individuals are theoretically able to maintain cooperation in a group by punishing noncooperators (4), we examined the response rates of each individual. However, because all subjects participated in four groups of differing composition, it was not practical to analyze the data statistically. Table S6 presents the data for each binary response (collapse or not collapse) for each individual and condition. There does not appear to be any individually consistent pattern of punishing third-party theft.

1. Forstmeier W, Schielzeth H (2011) Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behav Ecol Sociobiol* 65: 47–55.
2. Bolker BM, et al. (2009) Generalized linear mixed models: A practical guide for ecology and evolution. *Trends Ecol Evol* 24:127–135.
3. Jensen K, Call J, Tomasello M (2007) Chimpanzees are vengeful but not spiteful. *Proc Natl Acad Sci USA* 104:13046–13050.
4. O'Gorman R, Henrich J, Van Vugt M (2009) Constraining free riding in public goods games: Designated solitary punishers can sustain human cooperation. *Proc Biol Sci* 276: 323–329.

Table S1. Results of the likelihood ratio tests of full model vs. reduced model comparisons

	Likelihood ratio tests		
Factors included in the full model	χ^2	*df*	*P* value
Random effects factors			
Actor	12.584	1	0.0003*
Thief	0	1	1.0000
Victim	0.2747	1	0.6002
Control factors			
Session order effect	11.549	1	0.0007*
Trial order per session	0.3254	1	0.5684
Actor's age	2.7012	1	0.1003
Actor's sex	0.1619	1	0.6874
Actor's rearing history	1.4087	1	0.2353
Test factors			
Condition	19.508	5	0.0015*
Dominance of actor to thief	8.6729	1	0.0032*
Kinship of actor to victim	0.1476	1	0.7008

The sample comprises all actors (n = 13) and all test groups (n = 60). The full model was always compared with the reduced model not including the factor of interest. The table shows the χ^2, *df*, and *P* values of model comparison for each factor.
*Indicates a significant effect.

Table S2. Results of the likelihood ratio tests of model comparisons of the sample of all actors (n = 13) and all test groups (n = 60) in the 3P theft and 3P unfair condition

	Likelihood ratio tests		
Factors included in the full model	χ^2	*df*	*P* value
Random effects factors			
Actor	6.3079	1	0.0120
Thief	0	1	1.0000
Victim	0	1	1.0000
Control factors			
Session order effect	5.7068	1	0.0169*
Trial order per session	0.1261	1	0.7225
Actor's age	2.3232	1	0.1275
Actor's sex	0.3401	1	0.5598
Actor's rearing history	1.0868	1	0.2972
Test factors			
Condition	4.9506	2	0.08414′
Dominance of actor to thief	4.7679	3	0.1896
Kinship of actor to victim	2.4791	2	0.2895
Interaction of condition and dominance	0.1667	1	0.6831
Interaction of dominance and kinship	1.3138	1	0.2517

The full model was always compared with the reduced model not including the factor of interest. The table shows the χ^2, *df*, and *P* values of model comparison for each factor.
*Indicates a significant effect.

Table S3. Results of the final GLMM for all actors (13 actors, 60 groups)

Factor	Estimate	SE	*Z*	*P*
Dominance	0.6406	0.2221	2.884	0.0039*
Session order	−0.5875	0.1805	−3.254	0.0011*
Condition				
2P theft as reference				
2P theft – 2P loss	−0.5277	0.4989	−1.058	0.2902
2P theft – 3P theft	−0.7766	0.5174	−1.501	0.1334
2P theft – 3P loss	−0.5222	0.4976	−1.049	0.2939
2P theft – 3P unfair	−2.1749	0.6613	−3.289	0.0010*
2P theft – 3P no-victim	−1.9415	0.6232	−3.116	0.0018*
3P theft as reference				
3P theft – 2P loss	0.2489	0.5313	0.468	0.6395
3P theft – 2P theft	0.7766	0.5174	1.501	0.1334
3P theft – 3P loss	0.2544	0.5300	0.480	0.6312
3P theft – 3P unfair	−1.3983	0.6813	−2.052	0.0401*
3P theft – 3P no-victim	−1.1649	0.6440	−1.809	0.0705

The estimates display the regression weight of the factor on the response variable; SE is the SE of the estimate; the *Z* and *P* values are also given. The 2P theft and 3P theft were the reference conditions.
*Indicates a significant effect.

Table S4. Results of the GLMM for dominant actors (n = 12, groups = 30)

Factor	Estimate	SE	Z	P
Session order	−0.6530	0.2459	−2.655	0.0079*
Condition				
2P theft as reference				
2P theft – 2P loss	−1.6707	0.6699	−2.494	0.0126*
2P theft – 3P theft	−1.2475	0.6252	−1.995	0.0460*
2P theft – 3P loss	−1.7349	0.6748	−2.571	0.0101*
2P theft – 3P unfair	−2.8373	0.8855	−3.204	0.0013*
2P theft – 3P no-victim	−2.0360	0.7163	−2.842	0.0045*
3P theft as reference				
3P theft – 2P loss	−0.4232	0.7024	−0.602	0.5468
3P theft – 2P theft	1.2475	0.6252	1.995	0.0460*
3P theft – 3P loss	−0.4874	0.7038	−0.692	0.4886
3P theft – 3P unfair	−1.5899	0.9063	−1.754	0.0793
3P theft – 3P no-victim	−0.7886	0.7423	−1.062	0.2881

The estimates display the regression weight of the factor on the response variable; SE is the SE of the estimate; the Z and P values are also given. The 2P theft and 3P theft were the reference conditions.
*Indicates a significant effect.

Table S5. Results of the GLMM for subordinate actors (n = 12, groups = 30)

Factor	Estimate	SE	Z	P
Session order	−0.5989	0.3143	−1.905	0.0567
Condition				
2P theft – 2P loss	1.3948	0.8674	1.608	0.1078
2P theft – 3P theft	0.1775	0.9192	0.193	0.8468
2P theft – 3P loss	1.4752	0.8668	1.702	0.0887
2P theft – 3P unfair	−0.8714	1.0367	−0.841	0.4006
2P theft – 3P no-victim	−1.7458	1.2813	−1.363	0.1730
3P theft – 2P loss	1.2172	0.8785	1.386	0.1658
3P theft – 2P theft	−0.1775	0.9192	−0.193	0.8468
3P theft – 3P loss	1.2977	0.8775	1.479	0.1392
3P theft – 3P unfair	−1.0489	1.0489	−1.000	0.3173
3P theft – 3P no-victim	−1.9234	1.2918	−1.489	0.1365

The estimates display the regression weight of the factor on the response variable; SE is the SE of the estimate; the Z and P values are also given. The 2P theft and 3P theft were the reference conditions.
*Indicates a significant effect.

Table S6. Raw response data for each actor according to their rank relative to the thief

		2P theft		2P loss		3P theft		3P loss		3P unfair		3P no-victim	
Actor	No. of groups	Dom	Sub	Dom	Sub	Dom	Sub	Dom	Sub	Dom	Sub	Dom	Sub
Corry	5	0/1	0/4	0/1	0/4	0/1	0/4	0/1	0/4	0/1	0/4	0/1	0/4
Dorien	5	1/2	0/3	0/2	0/3	0/2	0/2*	0/2	0/3	0/2	0/3	0/2	0/3
Fraukje	4	1/3	0/1	1/3	0/1	2/3	0/1	1/3	0/1	0/3	0/1	1/3	0/1
Frodo	4	0/4	—	0/4	—	1/4	—	1/4	—	0/4	—	0/4	—
Lome	5	1/1	2/4	1/1	3/4	1/1	0/4	1/1	4/4	0/1	0/4	0/1	1/4
Natascha	5	2/3	0/1*	0/3	0/2	0/3	0/2	0/3	0/2	0/3	0/2	0/3	0/2
Patrick	5	2/3	0/2	2/3	1/2	2/3	1/2	0/3	1/2	1/3	1/2	1/3	0/2
Pia	5	0/1	0/4	0/1	0/4	0/1	0/3*	0/1	0/4	0/1	0/4	0/1	0/4
Riet	5	2/2*	0/2	0/3	0/2	0/3	0/2	1/3	0/2	1/3	0/2	1/3	0/2
Robert	4	0/4	—	0/4	—	1/4	—	0/4	—	0/4	—	1/4	—
Sandra	5	2/3	1/2	0/3	0/2	0/3	0/2	1/3	0/2	0/3	0/2	0/3	0/2
Tai	4	—	1/4	—	2/4	—	3/3*	—	3/4	—	1/4	—	0/4
Ulla	4	2/2	0/2	1/2	2/2	0/2	0/2	0/2	0/2	0/2	0/2	0/2	0/2

The number of groups the actor was tested in is also shown. The number before the slash is the number of trials in which the actor collapsed the trapdoor, and the number after the slash is the total number of trials for each condition. Dom, dominant; Sub, subordinate. *Indicates that there was one trial in which there was no theft (i.e., the individual in the thief's position did not pull food away from the actor); these cases were not included in the analysis, therefore the number of trials do not add up to the full complement. Dashes (—) indicate no data, namely that due to rank position, the subject was either never in the subordinate role (Frodo and Robert) or in the dominant role (Tai).

Table S7. Sex, age, and rearing histories of the chimpanzees tested

Name	Sex	Age (y)	Rearing history
Corry	Female	31	Nursery
Dorien	Female	28	Nursery
Fraukje	Female	32	Nursery
Frodo	Male	14	Mother
Lome	Male	7	Mother
Natascha	Female	28	Nursery
Patrick	Male	11	Mother
Pia	Female	9	Mother
Riet	Female	30	Nursery
Robert	Male	32	Nursery
Sandra	Female	15	Mother
Tai	Female	6	Mother
Ulla	Female	31	Nursery

Table S8. Number of groups tested

Dominance/kinship	Dominant	Subordinate
Kin	15	15
Nonkin	15	15

These groups are based on the actor's kinship with the victim and dominance to the thief.

Table S9. Details of the 60 test groups

Actor	Thief	Victim	Actor's dominance relationship with the thief	Actor's kin relationship to the victim	Test group
Corry	Frodo	Patrick	Subordinate	Mother	Subordinate/kin
Corry	Riet	Patrick	Subordinate	Mother	Subordinate/kin
Corry	Riet	Ulla	Subordinate	None	Subordinate/nonkin
Corry	Sandra	Lome	Dominant	Mother	Dominant/kin
Corry	Robert	Lome	Subordinate	Mother	Subordinate/kin
Dorien	Robert	Ulla	Subordinate	None	Subordinate/nonkin
Dorien	Lome	Corry	Dominant	None	Dominant/nonkin
Dorien	Riet	Robert	Subordinate	None	Subordinate/nonkin
Dorien	Fraukje	Patrick	Dominant	None	Dominant/nonkin
Dorien	Frodo	Corry	Subordinate	None	Subordinate/nonkin
Fraukje	Lome	Frodo	Dominant	None	Dominant/nonkin
Fraukje	Tai	Pia	Dominant	Mother	Dominant/kin
Fraukje	Sandra	Pia	Dominant	Mother	Dominant/kin
Fraukje	Frodo	Pia	Subordinate	Mother	Subordinate/kin
Frodo	Patrick	Natascha	Dominant	Infant	Dominant/kin
Frodo	Ulla	Patrick	Dominant	None	Dominant/nonkin
Frodo	Corry	Dorien	Dominant	None	Dominant/nonkin
Frodo	Pia	Riet	Dominant	None	Dominant/nonkin
Lome	Dorien	Corry	Subordinate	Infant	Subordinate/kin
Lome	Fraukje	Robert	Subordinate	None	Subordinate/nonkin
Lome	Tai	Pia	Dominant	None	Dominant/nonkin
Lome	Ulla	Patrick	Subordinate	Full-sibling	Subordinate/kin
Lome	Patrick	Natascha	Subordinate	None	Subordinate/nonkin
Natascha	Fraukje	Frodo	Dominant	Mother	Dominant/kin
Natascha	Pia	Dorien	Dominant	None	Dominant/nonkin
Natascha	Riet	Frodo	Subordinate	Mother	Subordinate/kin
Natascha	Robert	Corry	Subordinate	None	Subordinate/nonkin
Natascha	Tai	Frodo	Dominant	Mother	Dominant/kin
Patrick	Fraukje	Lome	Dominant	Full-sibling	Dominant/kin
Patrick	Sandra	Corry	Dominant	Infant	Dominant/kin
Patrick	Frodo	Lome	Subordinate	Full-sibling	Subordinate/kin
Patrick	Frodo	Corry	Subordinate	Infant	Subordinate/kin
Patrick	Tai	Lome	Dominant	Full-sibling	Dominant/kin
Pia	Natascha	Sandra	Subordinate	None	Subordinate/nonkin
Pia	Robert	Fraukje	Subordinate	Infant	Subordinate/kin
Pia	Patrick	Fraukje	Subordinate	Infant	Subordinate/kin
Pia	Corry	Lome	Subordinate	None	Subordinate/nonkin
Pia	Tai	Fraukje	Dominant	Infant	Dominant/kin
Riet	Ulla	Tai	Dominant	Mother	Dominant/kin
Riet	Lome	Sandra	Dominant	Mother	Dominant/kin
Riet	Robert	Dorien	Subordinate	None	Subordinate/nonkin
Riet	Patrick	Frodo	Subordinate	None	Subordinate/nonkin
Riet	Dorien	Sandra	Dominant	Mother	Dominant/kin
Robert	Dorien	Tai	Dominant	None	Dominant/nonkin
Robert	Sandra	Ulla	Dominant	None	Dominant/nonkin
Robert	Corry	Sandra	Dominant	None	Dominant/nonkin
Robert	Patrick	Riet	Dominant	None	Dominant/nonkin
Sandra	Tai	Robert	Dominant	None	Dominant/nonkin
Sandra	Natascha	Lome	Subordinate	None	Subordinate/nonkin
Sandra	Lome	Riet	Dominant	Infant	Dominant/kin
Sandra	Pia	Tai	Dominant	Full-sibling	Dominant/kin
Sandra	Frodo	Tai	Subordinate	Full-sibling	Subordinate/kin
Tai	Pia	Lome	Subordinate	None	Subordinate/nonkin
Tai	Frodo	Riet	Subordinate	Infant	Subordinate/kin
Tai	Lome	Sandra	Subordinate	Infant	Subordinate/kin
Tai	Natascha	Riet	Subordinate	Infant	Subordinate/kin
Ulla	Pia	Dorien	Dominant	None	Dominant/nonkin
Ulla	Frodo	Natascha	Subordinate	None	Subordinate/nonkin
Ulla	Riet	Lome	Subordinate	None	Subordinate/nonkin
Ulla	Tai	Fraukje	Dominant	None	Dominant/nonkin

Actors, thieves (when present), and victims (when present) constitute the groups. The dominance relationships between actor and thief, as well as kinship between actor and victim are shown. These combinations resulted in four test groups (dominant/kin, subordinate/kin, dominant/nonkin, and subordinate/nonkin). Because of possible agonistic interactions, the two adult males, Frodo and Robert, could not be tested together.