Universiteit
Leiden
The Netherlands

The (un)willingness to reward cooperation and punish non-cooperation

Molenmaker, Welmer E.

Citation
Molenmaker, W. E. (2017, January 19). *The (un)willingness to reward cooperation and punish non-cooperation*. *Kurt Lewin Institute Dissertation Series*. Retrieved from https://hdl.handle.net/1887/45536

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/45536

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/45536 holds various files of this Leiden University dissertation

Author: Molenmaker, Welmer E.
Title: The (un)willingness to reward cooperation and punish non-cooperation
Issue Date: 2017-01-19

Appendices

■ Appendix A

Supplemental experiment 1

Aim and Design

The aim of this pilot experiment was to investigate whether the anticipation of future interactions would affect the willingness to sanction in a public good dilemma. As third party, participants observed the choice behavior of two persons in a public good task, which either was the only round (one-shot conditions) or was the first of five rounds (multiple-shots conditions), and subsequently had the opportunity to administer increment coins (reward conditions) or decrement coins (punishment conditions). This experiment had a 2 (Sanction type: Punish versus Reward) × 2 (Interactions: One-shot versus Multiple-shots) design with Choice to sanction and Sanction size as dependent variables.

Results

The 2 (Sanction type) × 2 (Interactions) binary logistic regression on Sanction choice (N = 122) yielded only a significant Sanction type main effect (B = 1.39, SE = 0.51, Wald (df=1) = 7.33, p = .007, Odds Ratio = 3.40, 95% CI [1.47, 10.91]), which indicated that the proportion of participants choosing to punish (69.4%) was smaller than the proportion of participants choosing to reward (90%). The Interactions main effect (B = 0.32, SE = 0.47, Wald (df=1) = 0.48, p = .487, Odds Ratio = 1.38, CI [0.55, 3.45]) and the Sanction type × Interactions interaction effect (B = -0.62, SE = 1.06, Wald (df=1) = 0.34, p = .562, Odds Ratio = 1.85, CI [0.23, 14.84]) both were non-significant. The 2 (Sanction type) × 2 (Interactions) ANOVA on Sanction size (N = 122) yielded a significant Interactions main effect (F(1,118) = 3.94, p = .05, η^2 = .03, 90% CI [.00, .09]) and a significant Sanction type main effect (F(1,118) = 13.55, p < .001, η^2 = .10, CI [.03, .19]), which indicated that the size of the punishments (M = 16.45, SD = 24.34) was significantly smaller than the size of the rewards (M = 33.02, SD = 25.82). The Sanction type × Interactions interaction effect (F(1,118) = 0.25, p = .621, η^2 < .01, CI [.00, .04]) was non-significant.

Supplemental experiment 2

Aim and Design

The aim of this pilot experiment was to investigate the willingness to sanction in a common resource dilemma. As third party, participants observed the choice behavior of two persons in a one-shot common resource task, and subsequently had the opportunity to administer increment coins (reward conditions) or decrement coins (punishment conditions). This experiment had a One-factor (Sanction type: Punish versus Reward) design with Choice to sanction and Sanction size as dependent variables.

Results

The Chi-squared test on Choice to sanction (N = 83) showed that the proportion of participants choosing to punish (66.7%) was smaller than the proportion of participants choosing to reward

(85.4%), $\chi^2(1) = 3.97$, $p = .046$, Odds Ratio = 2.92, 95% CI [0.99, 8.57]. The One-way ANOVA on Sanction size ($N = 83$) showed that the size of the punishments ($M = 15.93$, $SD = 17.57$) was significantly smaller than the size of the rewards ($M = 26.29$, $SD = 21.79$), $t(81) = 2.39$, $p = .019$, $\eta^2 = .07$, 90% CI [.01, .17].

Supplemental experiment 3

Aim and Design

The aim of this experiment was to investigate whether advisors' lack of personal responsibility would affect their willingness to advice for sanctioning, as compared to administrators willingness to administer the sanctions. Participants observed the choice behavior of a group member in a common resource task, and subsequently had the opportunity to administer (administrator conditions) or advice an administrator about the administration of (advisor conditions) increment points (reward conditions) or decrement points (punishment conditions). This experiment had a 2 (Sanction type: Punish versus Reward) × 2 (Responsibility: Administrator versus Advisor) design with Choice to sanction and Sanction size as dependent variables.

Results

The 2 (Sanction type) × 2 (Responsibility) binary logistic regression on Sanction choice ($N = 157$) yielded a non-significant Sanction type main effect ($B = 0.17$, $SE = 0.55$, Wald (df=1) = 0.92, $p = .76$, Odds Ratio = 0.85, 95% CI [0.29, 2.49]), which indicated that the proportion of participants choosing to punish (91.1%) did not differ from the proportion of participants choosing to reward (89.7%). The Responsibility main effect ($B = 0.32$, $SE = 0.47$, Wald (df=1) = 0.48, $p = .487$, Odds Ratio = 1.38, CI [0.55, 3.45]) and the Sanction type × Responsibility interaction effect ($B = 0.28$, $SE = 1.11$, Wald (df=1) = 0.06, $p = .80$, Odds Ratio = 1.32, CI [0.15, 11.63]) were also non-significant. The 2 (Sanction type) × 2 (Responsibility) ANOVA on Sanction size ($N = 157$) yielded a non-significant Sanction type main effect ($F(1,153) = 1.08$, $p = .299$, $\eta^2 = .01$, CI [.00, .04]), which indicated that the size of the punishments ($M = 40.24$, $SD = 26.79$) did not differ from the size of the rewards ($M = 45.21$, $SD = 32.30$). The Responsibility main effect ($F(1,153) = 2.54$, $p = .113$, $\eta^2 = .02$, CI [.00, .06]) and Sanction type × Responsibility interaction effect ($F(1,153) = 0.01$, $p = .93$, $\eta^2 < .01$, CI [.00, .00]) were also non-significant.

Supplemental experiment 4

Aim and Design

The aim of this experiment was to investigate whether the timing of sanctioning decisions would affect the willingness to sanction when people themselves are involved in the public good dilemma. Participants observed the choice behavior of a group member in a public good task, and had the opportunity to administer increment points (reward conditions) or decrement points (punishment conditions), either before they received feedback (beforehand

conditions), directly after they received feedback (directly afterwards conditions), or after a time delay (delayed afterwards conditions). This experiment had a 2 (Sanction type: Punish versus Reward) × 3 (Decision timing: Beforehand versus Directly afterwards versus Delayed afterwards) design with Choice to sanction and Sanction size as dependent variables.

Results

The 2 (Sanction type) × 3 (Decision timing) binary logistic regression on Choice to sanction (N = 197) only yielded a significant Sanction type main effect (B = 1.74, SE = 0.44, Wald (df=1) = 15.84, p < .001, Odds Ratio = 5.68, 95% CI [2.41, 13.36]), while controlling for Donations (B = 0.08, SE = 0.06, Wald (df=1) = 1.57, p = .210, Odds Ratio = 1.08, CI [0.96, 1.22]), which indicated that the proportion of participants choosing to punish (68.4%) was smaller than the proportion of participants choosing to reward (91.9%). The Decision timing main effect (Wald (df=2) = 2.50, p = .287) and the Sanction type × Decision timing interaction effect (Wald (df=2) = 2.11, p = .348) were non-significant. The 2 (Sanction type) × 3 (Decision timing) ANOVA on Sanction size (N = 197) yielded a significant Decision timing main effect (F(1,190) = 3.96, p = .021, η^2 = .06, 90% CI [.01, .11]) and a significant Sanction type main effect (F(1,190) = 76.87, p < .001, η^2 = .28, CI [.19, .36]), while controlling for Donations (F(1,190) = 7.36, p = .007, η^2 = .03, CI [.00, .07]), which indicated that the size of the punishments (M = 22.04, SD = 27.67) was significantly smaller than the size of the rewards (M = 60.85, SD = 36.89). The Sanction type × Decision timing effect (F(1,190) = 0.10, p = .907, η^2 < .01, CI [.00, .01) was non-significant.

Supplemental experiment 5

Aim and Design

The aim of this pilot experiment was to investigate whether involvement in a public good dilemma would affect the willingness to sanction. Participants observed the choice behavior of a group member in a public good task, while they themselves either were involved in the task (second party conditions) or were not involved in the task (third party conditions), and subsequently had the opportunity to administer increment coins (reward conditions) or decrement coins (punishment conditions). This experiment had a 2 (Sanction type: Punish versus Reward) × 2 (Party type: Second party versus Third party) design with Choice to sanction and Sanction size as dependent variables.

Results

The 2 (Sanction type) × 2 (Party type) binary logistic regression on Choice to sanction (N = 156) yielded only a significant Sanction type main effect (B = 2.16, SE = 0.46, Wald (df=1) = 22.43, p < .001, Odds Ratio = 8.70, 95% CI [3.55, 21.29]), which indicated that the proportion of participants choosing to punish (53.9%) was smaller than the proportion of participants choosing to reward (91%). The Party type main effect (B = 0.78, SE = 0.39, Wald (df=1) = 0.04, p = .844, Odds Ratio = 1.08, CI [0.50, 2.34]) and the Sanction type

× Party type interaction effect (B = 0.32, SE = 1.92, Wald (df=1) = 0.12, p = .731, Odds Ratio = 1.37, CI [0.23, 8.32]) both were non-significant. The 2 (Sanction type) × 2 (Party type) ANOVA on Sanction size (N = 156) yielded only a significant Sanction type main effect (F(1,152) = 108.33, p < .001, η^2 = .41, 90% CI [.32, .49]), which indicated that the size of the punishments (M = 8.04, SD = 11.30) was significantly smaller than the size of the rewards (M = 49.28, SD = 33.11). The Party type main effect (F(1,152) = 0.42, p = .521, η^2 < .01, CI [.00, .03]) and the Sanction type × Party type effect (F(1,152) = 1.49, p = .224, η^2 = .01, CI [.00, .04) both were non-significant.

Supplemental experiment 6

Aim and Design

The aim of this pilot experiment was to investigate whether outcome dependence in a public good dilemma would affect the willingness to sanction. As third party, participants observed the choice behavior of group members in a public good task, while they either were dependent on the outcome of the task (third party dependence conditions) or were not dependent on the outcome of the task (third party independence conditions), and subsequently had the opportunity to administer increment coins (reward conditions) or decrement coins (punishment conditions). This experiment had a 2 (Party type: Third party dependence versus Third party independence) × 2 (Feedback: High cooperator versus Low cooperator) mixed design with repeated measures on the latter factor and Choice to reward or punish and Reward or punish size as dependent variables.

Results

The 2 (Party type) × 2 (Feedback) mixed binary logistic regression on Choice to reward or punish (N = 71) yielded only a significant Feedback main effect (B = -3.95, SE = 0.95, Wald (df=1) = 17.24, p < .001, Odds ratio = 7.46, 95% CI [3.36, 16.54]), which indicated that the proportion of participants choosing to punish the low cooperator (40%) was smaller than the proportion of participants choosing to reward the high cooperator (80%). The Party type main effect was non-significant (B = 0.71, SE = 0.50, Wald (df=1) = 2.00, p = .157, Odds ratio = 3.28, CI [0.62, 17.44]). The 2 (Party type) × 2 (Feedback) mixed ANOVA on Reward (positive value) or Punish (negative value) size (N = 75) yielded a marginal significant Party type main effect (F(1,53) = 3.60, p = .062, η^2 = .05, 90% CI [.00, .15]) and a significant Feedback main effect (F(1,73) = 43.61, p < .001, η^2 = .60, CI [.47, .68]), which indicated that the size of the punishments for the low cooperator (M = 16.71, SD = 28.06) was significantly smaller than the size of the rewards for the high cooperator (M = 42.68, SD = 36.49). The Party type × Feedback interaction effect (F(1,53) = 0.05, p = .821, η^2 < .01, CI [.00, .03]) was non-significant.

Supplemental experiment 7

Aim and Design

The aim of this experiment was to investigate whether involvement and outcome dependence in a public good dilemma would affect the willingness to sanction. Participants observed

the choice behavior of a group member in a public good task, while they themselves either were involved in the task (second party conditions), were dependent on the outcome of the task (third party dependence conditions), or were not dependent on the outcome of the task (third party independence conditions), and subsequently had the opportunity to administer increment coins (reward conditions) or decrement coins (punishment conditions). This experiment had a 2 (Sanction type: Punish versus Reward) × 3 (Party type: Second party versus Third party dependence versus Third party independence) design with Choice to sanction and Sanction size as dependent variables.

Results

The 2 (Sanction type) × 3 (Party type) binary logistic regression on Choice to sanction (N = 284) only yielded a significant Sanction type main effect (B = 1.48, SE = 0.31, Wald (df=1) = 22.42, p < .001, Odds Ratio = 4.40, 95% CI [2.38, 8.12]), which indicated that the proportion of participants choosing to punish (62.5%) was smaller than the proportion of participants choosing to reward (97.9%). The Party type main effect (Wald (df=2) = 3.36, p = .186) and the Sanction type × Party type interaction effect (Wald (df=2) = 0.52, p = .771) were non-significant. The 2 (Sanction type) × 3 (Party type) ANOVA on Sanction size (N = 284) yielded a significant Party type main effect (F(1,278) = 3.44, p = .033, η^2 = .04, 90% CI [.01, .08]) and a significant Sanction type main effect (F(1,278) = 47.14, p < .001, η^2 = .14, CI [.08, .20]), which indicated that the size of the punishments (M = 29.99, SD = 33.51) was significantly smaller than the size of the rewards (M = 56.79, SD = 32.45). The Sanction type × Party type effect (F(1,278) = 0.90, p = .407, η^2 = .01, CI [.00, .03) was non-significant.

APP

Supplemental experiment 8

Aim and Design

The aim of this pilot experiment was to investigate whether the experience of shame would affect the willingness to sanction in a public good dilemma. After a shame inducing task, participants observed the choice behavior of a group member in a public good task and subsequently had the opportunity to administer increment coins (reward conditions) or decrement coins (punishment conditions). This experiment had a 2 (Sanction type: Punish versus Reward) × 2 (Shame versus Control) design with Choice to sanction and Sanction size as dependent variables.

Results

The 2 (Sanction type) × 2 (Emotion condition) binary logistic regression on Choice to sanction (N = 148) yielded only a marginal significant Sanction type main effect (B = 0.86, SE = 0.45, Wald (df=1) = 3.60, p = .058, Odds Ratio = 2.35, 95% CI [0.97, 5.69]), which indicated that the proportion of participants choosing to punish (75.7%) was smaller than the proportion of participants choosing to reward (87.8%). The Emotion condition main effect (B = 0.66, SE = 0.44, Wald (df=1) = 2.23, p = .135, Odds Ratio = 1.94, CI [0.81, 4.63]) and the Sanction type × Emotion condition interaction effect (B = 1.12, SE = 1.00,

Wald (df=1) = 1.24, p = .266, Odds Ratio = 3.04, CI [0.43, 21.62]) both were non-significant. The 2 (Sanction type) × 2 (Emotion condition) ANOVA on Sanction size (N = 148) yielded only a significant Sanction type main effect (F(1,144) = 36.40, p < .001, η^2 = .20, 90% CI [.11, .29]), which indicated that the size of the punishments (M = 1.68, SD = 1.73) was significantly smaller than the size of the rewards (M = 4.11, SD = 2.98). The Emotion condition main effect (F(1,144) = 0.22, p = .64, η^2 < .01, CI [.00, .03]) and the Sanction type × Emotion condition effect (F(1,144) < 0.01, p = .947, η^2 < .01, CI [.00, .00) both were non-significant.

Supplemental experiment 9

Aim and Design

The aim of this pilot experiment was to investigate whether the experience of guilt would affect the willingness to sanction in a public good dilemma. After a guilt inducing task, participants observed the choice behavior of a group member in a public good task and subsequently had the opportunity to administer increment coins (reward conditions) or decrement coins (punishment conditions). This experiment had a 2 (Sanction type: Punish versus Reward) × 2 (Guilt versus Control) design with Choice to sanction and Sanction size as dependent variables.

Results

The 2 (Sanction type) × 2 (Emotion condition) binary logistic regression on Choice to sanction (N = 147) yielded only a significant Sanction type main effect (B = 1.74, SE = 0.49, Wald (df=1) = 12.51, p < .001, Odds Ratio = 5.70, 95% CI [2.17, 14.94]), which indicated that the proportion of participants choosing to punish (66.2%) was smaller than the proportion of participants choosing to reward (91.8%). The Emotion condition main effect (B = -0.01, SE = 0.43, Wald (df=1) = 0.00, p = .987, Odds Ratio = .993, CI [0.43, 2.29]) and the Sanction type × Emotion condition interaction effect (B = 0.87, SE = 1.03, Wald (df=1) = 0.71, p = .399, Odds Ratio = 2.38, CI [0.32, 17.75]) both were non-significant. The 2 (Sanction type) × 2 (Emotion condition) ANOVA on Sanction size (N = 147) yielded a significant Emotion condition main effect (F(1,143) = 5.37, p = .022, η^2 = .03, 90% CI [.00, .09]), a significant Sanction type × Emotion condition effect (F(1,143) = 4.78, p = .03, η^2 = .03, CI [.00, .08), and a significant Sanction type main effect (F(1,143) = 30.96, p < .001, η^2 = .09, CI [.08, .26]), which indicated that the size of the punishments (M = 2.00, SD = 2.26) was significantly smaller than the size of the rewards (M = 4.26, SD = 2.85).

■ Appendix B

Two separate meta-analyses were conducted to estimate the combined overall effect of Sanction type (Reward versus Punishment) on Choice to sanction and on Sanction size. The data were taken from experiments reported in the empirical chapters of this present dissertation (Chapters 2-4) and from experiments – conducted by the author – not included in the empirical chapters of this present dissertation (Appendix A). The inclusion criteria's were that (1) sanction type was manipulated between participants – thereby excluding Experiment 2.2, Experiment 3.3, and Supplemental Experiment 6 – and that (2) Choice to sanction (k = 13, n = 2073) and/or Sanction size (k = 13, n = 2056) was measured.

Meta-analytic procedures

For the meta-analysis on Choice to sanction, the Odds ratio statistic was used as measure of effect size. The Odds ratios were calculated in the *Meta-Essentials* workbook using the frequencies of participants choosing to sanction along with the cell sizes (Van Rhee, Suurmond, & Hak, 2015). For the meta-analysis on Sanction size, the Cohen's *d* statistic was used as measure of effect size. The Cohen's *d*'s were calculated in the *Meta-Essentials* workbook using the *F* score or *t* value along with the cell sizes (Van Rhee et al., 2015).

Since sanction type was in most of the studies not the only factor that was manipulated between participants, it was assumed that the effect of Sanction type on Choice to sanction and Sanction size will have systematic variation (i.e., heterogeneity). To estimate the average effect sizes for Choice to sanction and Sanction size, random-effects models were therefore used since these models assume that effect sizes are sampled from a population of varying effect sizes (e.g., Hedges & Vevea, 1998). Finally, the meta-analyses were conducted using the inverse variance weighting method, both in the *Meta Essentials* workbooks (Van Rhee et al., 2015).

Results

Choice to sanction

The effect size estimations and their 95% confidence intervals of the experiments used in the meta-analysis on Choice to sanction (k = 13) are reported in Table 6.1. In accordance with my prediction, Sanction type had a medium-sized overall effect (see also Chen, Cohen, & Chen, 2010) on Choice to sanction (Z = 6.24, p < .001, Odds ratio = 3.39, 95% CI [2.22, 5.19]), which indicated that non-cooperation was punished less often than cooperation was rewarded. As expected, there is heterogeneity in the distribution of effect sizes (T^2 = .27, T = .52, I^2 = 58.46%). Thus, the combined overall effect of Sanction type on Choice to sanction should not be treated as the 'true' effect size (see Hak, Van Rhee, & Suurmond, 2016).

Sanction size

The effect size estimations and their 95% confidence intervals of the experiments used in the meta-analysis on Sanction size (k = 13) are reported in Table 6.2. In accordance with my

Table 6.1. *Odds ratios and their 95% Confidence Intervals per Experiment*

	Odds Ratio	95% Confidence Intervals	
Experiment 2.1	4.98	1.83	13.59
Experiment 3.1	1.12	0.57	2.21
Experiment 4.1	2.31	1.07	4.96
Experiment 4.2	3.67	1.73	7.80
Experiment 4.3 – Footnote	11.77	2.53	54.84
Supplemental Experiment 1	3.98	1.45	10.94
Supplemental Experiment 2	2.92	0.98	8.71
Supplemental Experiment 3	0.85	0.29	2.49
Supplemental Experiment 4	5.26	2.26	12.24
Supplemental Experiment 5	8.69	3.53	21.43
Supplemental Experiment 7	4.34	2.35	8.00
Supplemental Experiment 8	2.32	0.96	5.62
Supplemental Experiment 9	5.70	2.15	15.06

prediction, Sanction type had a large-sized overall effect (see also Cohen, 1988) on Sanction size ($Z = 8.79$, $p < .001$, $d = 0.94$, 95% CI [0.71, 1.18]), which indicated that non-cooperation was punished to a lesser extent than cooperation was rewarded. As expected, there is heterogeneity in the distribution of effect sizes ($T^2 = .11$, $T = .33$, $I^2 = 79.35\%$). Thus, the combined overall effect of Sanction type on Sanction size should not be treated as the 'true' effect size (see Hak et al., 2016).

Table 6.2. *Cohen's d's and their 95% Confidence Intervals per Experiment*

	Cohen's *d*	95% Confidence Intervals	
Experiment 2.1	1.00	0.62	1.39
Experiment 3.2	0.73	0.41	1.06
Experiment 4.1	1.15	0.82	1.49
Experiment 4.2	0.98	0.69	1.26
Experiment 4.3 – Footnote	1.44	1.02	1.86
Supplemental Experiment 1	0.67	0.30	1.03
Supplemental Experiment 2	0.52	0.08	0.97
Supplemental Experiment 3	0.17	-0.15	0.48
Supplemental Experiment 4	1.25	0.94	1.56
Supplemental Experiment 5	1.67	1.30	2.03
Supplemental Experiment 7	0.81	0.57	1.06
Supplemental Experiment 8	0.99	0.65	1.34
Supplemental Experiment 9	0.92	0.58	1.26