Punishment, reward, and cooperation in a framed field experiment

Noussair, Charles and van Soest, Daan and Stoop, Jan

Tilburg University, CentER, VU University Amsterdam, Erasmus University Rotterdam

26 September 2011

Punishment, Reward, and Cooperation in a Framed Field Experiment*

Charles N Noussair[†]
Daan van Soest[‡]
Jan Stoop[§]

October 12, 2011

Abstract

We report a framed field experiment, in which we study the effectiveness of punishment and reward in sustaining cooperation in a social dilemma. Punishments tend to be directed at non-cooperators and rewards are assigned by those who are relatively cooperative. In contrast to the results typically found in laboratory experiments, however, we find that punishments and rewards fail to increase cooperation.

JEL Classification: C72, C92, C93.

Keywords: Field experiment, public goods game, social preferences, punishment, reward

*We are grateful to the Netherlands Organization for Scientific Research, NWO, for financial support as part of the Program on Evolution and Behavior. We would like to thank Ben and Shirley Willems for use of their fishing facility, and Paul Ludeña Delgado, Joris Hoendervangers, Stef van Kessel, and Menusch Khadjavi for excellent research assistance. We thank seminar participants at Iowa State and Nottingham Universities for helpful comments.

†Tilburg University, Department of Economics and CentER, Tilburg, the Netherlands.

‡Free University Amsterdam, Department of Spatial Economics and IVM, Amsterdam, The Netherlands, and Tilburg University, Department of Economics and CentER, Tilburg, the Netherlands.

§Erasmus University Rotterdam, Department of Applied Economics, Rotterdam, the Netherlands. Phone: 003110-89517, e-mail: stoop@ese.eur.nl, adres: Burgemeester Oudlaan 50, PO Box 1738, 3000 DR Rotterdam, The Netherlands

1 Introduction

The search for factors that promote cooperation in social dilemmas is a focus of research in the social sciences (see for example Ostrom (1990)). Controlled laboratory experiments show that cooperation among members of a group can be promoted with a system of decentralized peer punishment (see for example Yamagishi (1986), Ostrom et al. (1992), Fehr and Gächter (2000), Fehr and Gächter (2002))), provided that the cost to the sanctioner is sufficiently low (Anderson and Putterman (2006), Nikiforakis and Normann (2008), Carpenter (2007)). Similarly, if it is not too costly to do so, the ability to reward peers can also increase cooperation (Sefton et al. (2007), Vyrastekova and van Soest (2008), and Rand et al. (2009)). The individual-level patterns of punishment and reward in the laboratory have the intuitive features that it is typically non-cooperators who are punished and cooperators who receive rewards. Non-cooperators respond on average to receiving punishment by increasing their level of cooperation.

The generality of the ability of punishment to promote cooperation in a social dilemma has been called into question in a recent paper by Janssen et al. (2010). They consider a laboratory environment, which includes spatial and dynamic elements that exist in many of the social dilemmas found in the field, in particular those involving extraction of a common pool resource. Their environment has a resource located throughout a grid, and group members can observe each other and adjust their behavior in real time in response to others' pace of extraction. The authors find that punishment is ineffective in reducing extraction of the resource, and thus fails to promote cooperation, unless it is accompanied by the ability of players to communicate with each other. They conjecture that the relative complexity of the environment (by experimental standards) makes punishments more difficult to interpret for the recipient. They write "... in a modestly complex dynamic and spatial environment where participants can punish back but cannot discuss why they are punished, receiving a sanction does not carry a clear message. Does the sanction relate to the amount harvested, the location, the spatial pattern of harvesting, the speed in which the avatar moves over the screen, etc. . . ?"

In this paper, we consider the effectiveness of punishment and reward in promoting cooperation using a controlled field experiment. One typical form of field experimentation is to alter one or more features of a laboratory paradigm to conform more closely to field conditions, but in a manner that allows the researcher to make causal inferences. This typically comes at a cost of some degree of control of the experiment. Usually, the differences between laboratory and field experiments lie in the context in which subjects make decisions, the stakes at play, or the characteristics of participants (see Harrison and List (2004) for a detailed discussion). The extent to which laboratory results concerning pro-social behavior carry over to non-laboratory settings has been the subject of recent debate (see Levitt and List (2007), Levitt and List (2008), and Falk and Heckman (2009)).

The setting of our experiment is a recreational fishing facility. The participants are regular customers of the facility. The fishermen are assigned to groups of four members, and play a game with incentives very similar to the linear Voluntary Contributions game (see Ledyard (1995) or Chaudhuri (2010)), though with arguably a framing closer to a common pool resource environment. The interaction is structured so that catching fewer fish gives rise to a positive externality on other group members by increasing the time that these others are allowed to fish. The incentives ensure that a social dilemma exists. There are three treatments. In the *Punishment* treatment, individuals have an opportunity to punish others after observing how many fish they catch. In the *Reward* treatment, they can reward others. Punishment (reward) takes the form of decreases (increases) in the recipient's fishing time. In the *Baseline* treatment, they can neither punish nor reward. The experiment has many of the features of the Janssen et al. (2010) paradigm. In particular, individuals choose a cooperation intensity in real time, they can update it in response to the real-time behavior of others, and the setting is framed as an extraction of a resource. The design of the experiment is described in section two.

In contrast to results reported from most laboratory experiments, but in agreement with Janssen et al. (2010), we find no evidence that punishment promotes cooperation. Rewards are also ineffective. Fishermen fish with the same intensity in the Baseline, Punishment and Reward treatments. Indeed, there is no evidence of cooperation in any of the treatments. Nevertheless, punishments, when imposed, are targeted at non-cooperators. Rewards are used more often than punishments. Reward assignments are uncorrelated with the cooperation level of the recipient, but are generally awarded by those who cooperate in the social dilemma. Thus, punishment assignments are dependent on the cooperation decisions of the recipient, while reward decisions actions depend on those of the allocator.

Our results support three more general contentions that previous authors have advanced. The first is that peer-to-peer punishment is not necessarily effective in promoting cooperation, suggesting that the findings of Janssen et al. (2010) are not an anomaly. The similarities between our field setting and their laboratory environment suggest that it is features of the game itself, rather than the fact that our study is a field experiment, which account for the ineffectiveness of punishment. The second is that our results are consistent with the view that punishment is not necessarily applied instrumentally to increase cooperation. Because it is ineffective in increasing cooperation, and is assigned by individuals who are not themselves cooperating more in response to receiving punishment, it is likely that some punishment is assigned with no expectation of modifying the subsequent behavior of the recipient. This pattern is consistent with the notion that individuals have a taste for punishing non-cooperators (Fehr and Gächter (2000), de Quervain et al. (2004), Singer et al. (2006)). The third point is that rewards can also be ineffective in promoting cooperation.

The remainder of this paper is organized as follows. In section 2, the setting, the experimental design and the procedures are described. Section 3 presents the data analysis, and section 4 concludes.

2 The setting, experimental design and experimental procedure

2.1 The setting of the field experiment

The experiment was conducted at a privately owned recreational fishing facility called 'De Biestse Oevers'.[1] The site consists of three outdoor rectangular ponds, each of which is roughly 3500 square feet in area. Each pond has room for twenty fishermen, ten on the the east and ten on the west side. A half-day of fishing costs either €12.50 or €15. For €12.50, four rainbow trout are put into the pond while for €15, two rainbow trout and one salmon trout, a larger trout species, are put in. Depending on the season, a customer can either fish for four or five hours. A fisherman is allowed to catch as many fish as possible within this time frame, but no compensation is provided if a fisherman does not catch the fish that are released on his behalf. There are strict rules in effect governing how individuals may fish. For example, it is prohibited to use more than one rod at the same time, no scoop net can be used to catch fish, and each fish caught has to be taken away from the site when the fisherman leaves.

The setting has a number of features conducive to its use for experimentation. First, the setting allows the experimenter to retain a strong degree of control, since the pond used is self-contained, dedicated exclusively to the experiment while a session is in progress, communication between participants can be monitored and restricted, and precise measures of cooperation can be gathered. Second, the participant pool of recreational fishermen is experienced with imposing negative externalities on one another. Because the number of fish in the pond is fixed at the start of a regular fishing day (when no experiment is being conducted), each fish a fisherman catches reduces the quantity of fish available to others. Third, individuals sampled from the same subject pool are known to have behaved similarly to individuals drawn from a pool of undergraduate students, in an abstract laboratory social dilemma conducted for another study (Stoop et al. (2010)). That is,

[1] For photos, we refer to www.biestse-oevers.nl.

in the laboratory experiment, the fishermen exhibit positive but less than full cooperation, which then declines over time (see Ledyard (1995). This indicates that any differences between our experiment and results from the laboratory are unlikely to be due purely to subject pool differences. Fourth, the context in which the current field experiment is implemented is natural and familiar to the participants. The salient rewards in the field experiment are fish and fishing time, the items that typically have value at the pond. This makes it easier for participants to understand the rules of the game.

As indicated earlier, there were three treatments, called Baseline, Punishment, and Reward. The next subsection describes the Baseline treatment, while the other two conditions are presented in section 2.3.

2.2 The Design of the Baseline treatment

In the Baseline treatment, participants are placed in groups of four. Group membership is fixed for the entire session, but no participant is informed about the identity of the three other members of her group. A session of the Baseline treatment consists of two parts. Part 1 consists of three periods of thirty minutes each, during which each participant is allowed to catch at most two fish per period. Part 2 lasts for at most 150 minutes, during which each participant is allowed to catch as many fish as he can. However, the actual duration of part 2 for each individual participant depends on how many fish each of the three other members of his group catch in part 1.

At the beginning of part 1, two rainbow trout per participant are put into the pond, plus an additional six trout. In each session, 16 fishermen participate, and hence we always throw in 38 fish at the start of a session. The spot at which a fisherman fishes, is assigned by a lottery. In each of the three thirty-minute periods, each fisherman is allowed to catch up to two fish. After he catches a second fish within a period, the fisherman has to stop fishing and wait until the start of the next period. The fisherman is allowed to keep all fish he catches, and he does not receive any monetary payment in part 1.

At the end of each period within part 1, each fisherman is informed of the number of fish caught by each of the other members of his group in the period. This information is presented on a sheet of paper in a new random order in each period. This ensures that participants cannot link the catch of one group member in the current period to the number of fish he caught in prior periods. At the start of each subsequent period, the number of fish caught in the previous period is replenished. This means that, at the start of each period within a session, an equal number of trout is present in the pond. Thus, at least in principle, it is equally easy to catch a fish in each period within part 1. In part 2, participants are free to catch as many fish as they would like. In addition to being allowed to keep the fish, individuals receive a bonus of €2 for each fish that they catch.

The social dilemma is introduced by the imposition of the following rule. Each fish that a fisherman catches in part 1 reduces the length of time that each of the other three members of his group can fish in part 2, by ten minutes. Assuming that a fisherman has monotonically increasing preferences over fishing time and money, fishing in part 2 is more valuable than in part 1. This is because (i) each fisherman is not constrained in the number of fish that he is allowed to catch, and (ii) each fish caught yields a monetary bonus of €2, while no such bonus was paid in part 1. We keep the level of difficulty of catching fish roughly constant within part 2 by replenishing the stock of fish every 30 minutes, throwing in the same number of fish as were caught in the preceding 30 minutes.

Thus, the social optimum is attained if all group members catch no fish in part 1. However, the subgame perfect Nash equilibrium strategy for each fisherman is to fish at full force in each period of part 1, since only other group members are harmed if an individual catches a fish. If all participants catch their maximum allowable quantity of two fish per period in each of the three periods of part 1, each participant's available fishing time in part 2 is reduced from 150 to 0 minutes.[2]

To verify that participants perceived the game as a social dilemma, we

[2]Because of the non-negativity constraint on time, a participant's fishing time in part 2 is set to zero if the other three participants in his group catch more than fifteen fish in part 1.

surveyed regular customers of the fishing pond, asking them which of the following options they valued more highly: (i) to wait for ninety minutes and then have 150 minutes of unconstrained fishing while receiving €2 per fish caught, or (ii) to fish for ninety minutes with the right to catch up to six fish. Option (i) is the strategy profile that is meant to correspond to full cooperation, and (ii) is the profile corresponding to selfish behavior. Of the 21 fishermen we surveyed, 19 indicated a preference for option (i) over option (ii), and hence we are confident that indeed our parametrization induces a social dilemma. The text and results of the questionnaire are reprinted in Appendix B.

2.3 The Punishment and Reward treatments

All of the rules described in the previous subsection for the Baseline treatment are also in effect in both the Punishment and Reward treatments. However, in the latter two treatments, subjects are also given an opportunity to directly reciprocate to the behavior of the other three members of their group. At the end of each of the three periods within part 1, subjects receive information about the catch of each other group member. In the Punishment (Reward) treatment, subjects can then choose to give up minutes of their own part 2 fishing time in order to reduce (increase) the part 2 fishing time of a designated group member, by three times as many minutes. Punishments or rewards can be imposed in blocks of five minutes (each reducing or increasing the recipient's part 2 fishing time by fifteen minutes). Subjects cannot spend more than a total of three blocks of five minutes on punishments or rewards in a single period. All punishment or reward decisions are made without knowing others' reward or punishment decisions for the current period.

If a subject's period 2 fishing time has been reduced to zero as a result of the fishing behavior of other members of his group and/or his own use of punishment or reward blocks in part 1, he can not assign any punishments or rewards. A fisherman could never have a balance of less than zero minutes of part 2 fishing time. In the Reward treatment, subjects could end up with more than 150 minutes fishing time in part 2. In the event that their time extended beyond the closing time of the facility for the day, subjects would receive coupons that could be used at the facility at any future date to reduce the entrance fee that is normally charged. Each coupon was worth €3, and participants would receive one for every thirty minutes of unused part 2 fishing time remaining.

After each subject has made his punishment or reward decision for the current period, he is informed about (i) the total number of five minute blocks (of rewards or punishments, depending on the treatment) the other three group members have assigned to him, (ii) the total amount of part 2 fishing time he currently has available, and (iii) the total number of five minute blocks of either rewards or punishments that were assigned to each of the other members of his group.

2.4 Measurement of cooperation

Measuring cooperation in our field setting is less straightforward than in laboratory experiments. In principle, cooperation may be measured by the amount of allowable catch not caught by a group, as reflected by the following equation:

$$C = \sum_i x_{it}^{Baseline}/n - \sum_i x_{it}^{\{Pun,Rew\}^j}, \qquad (1)$$

where $\sum_i x_{it}^{\{Pun,Rew\}^j}$ is the total catch in period t of group j in either the Punishment or Reward treatment, and $\sum_i x_{it}^{Baseline}/n$ is the average catch in period t of all n groups in the Baseline treatment. A value of C equal to 0 would indicate the absence of a treatment effect, and a positive level of C would indicate the presence of additional cooperation in the Punishment or Reward treatments relative to the Baseline. We will refer to this measure as the *Catch* measure of cooperation.

An advantage of this measure is that it is easy to observe. A disadvantage is that subjects may not always be able to catch all fish they would like to, because of insufficient skill or because of exogenous factors such as weather conditions. Therefore, measuring cooperation by the remaining quantity of allowable catch may overstate the extent to which subjects act cooperatively. Fortunately, the setting allows us to employ a second measure of cooperation based on the amount of effort subjects put into catching fish. Rainbow trout are predators that actively pursue prey. When bait is dragged through the water it naturally attracts the fish and the more the bait is exposed to trout, the higher is the probability of catching a fish. Therefore, a fisherman can increase his expected catch by casting his bait more frequently. Indeed, in appendix A we show that there is a positive correlation between effort and catch; the more often a fisherman casts his rod, the more fish he catches. To gather data on how frequently participants cast their fishing rods, two experimenters monitored the sixteen fishermen that were present in any session and counted the number of times they cast their rods in every 5 minute time interval.

As a normalized measure of cooperation, we report the number of times a fisherman casts his bait per minute. A comparison of effort levels in the Punishment and Reward treatments, relative to Baseline, indicates whether the instruments increase cooperation. We use an analogous measure of cooperation to that presented in equation (1), but replacing the quantity of fish caught per period with the average number of casts per minute. We term the resulting value the *Effort* measure of cooperation.

2.5 Experimental procedures

The experiments were conducted in April and May of 2010. Recruitment was done two weeks in advance by handing out flyers at the fishing site. All sessions had sixteen participants, and no fisherman participated more than once. Upon arrival at the pond, the fishing spots for part 1 of the experiment were randomly assigned. The participants received instructions in groups of four at each of the four corners of the pond. Each participant in an instruction group was subsequently reassigned to a different group for the fishing activity, and participants were informed about this on beforehand. The instructions were given in this manner to prevent communication between members of the same group. The instructions were read out aloud by the experimenters, and the participants received handouts with the summary of the instructions, which they could keep for the remainder of the experiment. After the instructions, each fisherman individually answered some control questions. The experiment began only once all participants had correctly answered the test questions.

At the end of each period of part 1, an experimenter informed the participants about the number of fish each of the other three participants in their group had caught in that period. Catch data were presented anonymously and in a random order. This rendered participants unable to link individual catch data between periods, from which they might be able to infer the identity of their fellow group members. Each participant was also informed about his amount of part 2 fishing time remaining. In the Punishment and Reward treatments, players were also informed of the punishment assigned to them by each other member of their group, in a format which ensured that the identity of the assigning players remained anonymous.

Table 1 presents the number of groups that participated in each treatment, and also the number of participants. In the analysis that follows, each group's average activity over a session is taken as the independent unit of observation.

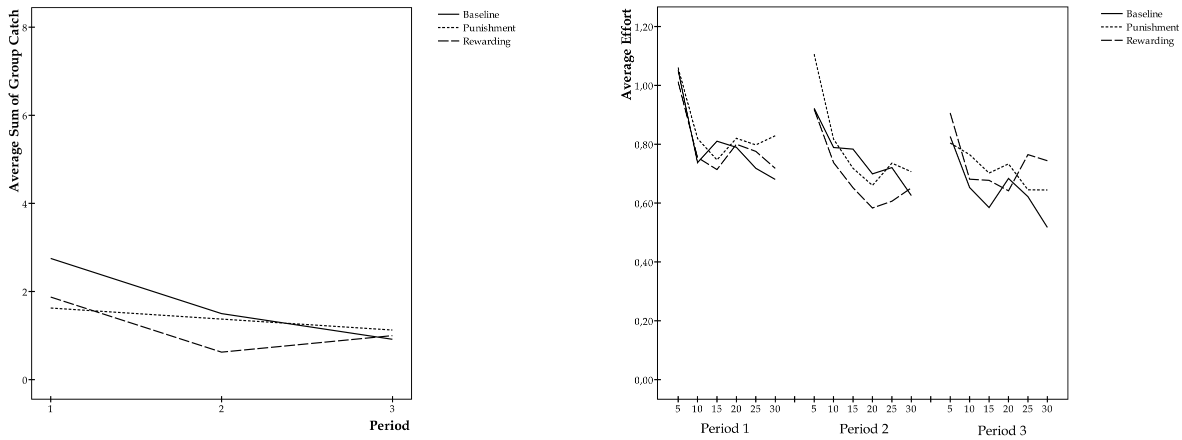
Treatment	Number of Groups	Number of participants
Baseline	12	48
Punishment	8	32
Reward	8	32

Table 1 Summary information about the three treatments.

3 Results

3.1 Cooperation

Punishment and reward opportunities have no impact on cooperation in our data. This is the case independently of whether the Catch or the Effort measure of cooperation is employed. Figure 1 shows the average number of fish caught per group (panel a) as well as the group averages of fishing effort (panel b).



(a) (b)

Figure 1 **(a)** Average number of fish caught per group in the Baseline, Punishment, and Reward treatments. **(b)** Average effort per group in the three treatments presented per five minute interval.

Figure 1 shows that the absolute levels of catch and effort in the Baseline, Punishment and Reward treatments are very similar over the three periods of part 1. There is a decline in the number of fish caught, as well as in the level of effort, over time. However, this decline in catch and effort is similar in all treatments, which implies that neither punishment nor reward has any effect on cooperation. This finding constitutes our first result:

Result 1 The availability of punishment or reward does not change the level of cooperation. Average Catch and Effort are the same in Part 1 in all three treatments.

Support for result 1: We perform a series of Mann-Whitney tests. As an independent observation, we use the average catch or effort level of a group during part 1. The catch in the Punishment treatment does not differ from the catch in the Baseline treatment ($N_1 = 12, N_2 = 8, p = 0.473$). Likewise, there is no significant difference between catch in the Baseline and the Reward treatments ($N_1 = 12, N_2 = 8, p = 0.427$). There is also no significant difference in catch between the Punishment and Reward treatments ($N_1 = 8, N_2 = 8, p = 0.195$). With regard to the effort levels, there is no significant difference between the Punishment and the Baseline treatments ($N_1 = 12, N_2 = 8, p = 0.521$), between the Reward and the Baseline treatments ($N_1 = 12, N_2 = 8, p = 0.851$), or between the Punishment and the Reward treatments ($N_1 = 8, N_2 = 8, p = 0.798$). ■

The data presented in Figure 1 illustrate the lack of treatment effects. We now consider whether cooperation occurs at all. We do this in two ways. Evidence of partial cooperation can appear in patterns within or between periods. Players may reduce their effort (possibly to zero) to catch a second fish in a period after they have caught the first fish in that period. This would reveal a restraint in effort in catching the second fish, since the fisherman is exerting less effort than he is able to. Players may also exhibit less effort in part 1 than in part 2. This would show cooperation since only in the former does catching a fish induce a negative externality on other players. We first consider within-period cooperation, and report our finding as result 2.

Result 2 There is no evidence of partial cooperation within a period in any of the three treatments.

Support for result 2: We compare the effort to catch the first and second fish within a period. A series of Wilcoxon matched pairs tests shows that effort levels remain the same after the first fish is caught. Each observation for these tests is the effort level of one fisherman in one period, in which he has caught at least one fish. The matched pair consists of the effort level of the individual during the time when he was trying to catch the first fish in a period, and the effort level while trying to catch the second fish. In the Baseline treatment, the effort levels for the first fish are not significantly different from those to catch the second fish in period 1 ($N_1 = N_2 = 24, p = 0.768$), in period 2 ($N_1 = N_2 = 15, p = 0.198$), or in period 3 ($N_1 = N_2 = 9, p = 0.401$). Likewise, in the punishment treatment, the levels of effort to catch the second fish are equal to those to catch the first fish in period 1 ($N_1 = N_2 = 9, p = 0.594$), period 2 ($N_1 = N_2 = 8, p = 1.000$) and period 3 ($N_1 = N_2 = 7, p = 0.310$). The same holds for the Reward treatment, where the effort levels to catch the second fish are similar to those to catch the first fish in period 1 ($N_1 = N_2 = 10, p = 0.575$), period 2 ($N_1 = N_2 = 4, p = 0.273$) and period 3 ($N_1 = N_2 = 6, p = 0.345$). ■

As the second test for the existence of partial cooperation, we compare catch and effort between parts 1 and 2 of the experiment. In part 2, there are no negative externalities and there is even an additional incentive to try to catch more fish (the €2 bonus per fish caught). However, when comparing effort in part 2 to that in period 3 of part 1, we actually find that effort is greater in part 1.

Result 3 Average effort in the last period of part 1 is greater than that in part 2 in all treatments.

Support for result 3: We use a Wilcoxon matched pairs test, taking the average effort level of a group in period 3, the last period, of part 1, and the average effort level of the same group in part 2, as an independent pair. Effort levels in period 3 of part 1 are greater in the Baseline treatment ($N_1 = N_2 = 12, p = 0.002$), the Punishment treatment ($N_1 = N_2 = 8, p = 0.024$), and the Reward treatment ($N_1 = N_2 = 12, p = 0.018$). ■

3.2 Punishment and reward assignments

We now analyze patterns in punishment and reward assignments. Figure 2 shows the average number of five minute blocks distributed per group over the three periods of part 1 in the Punishment and Reward treatments.

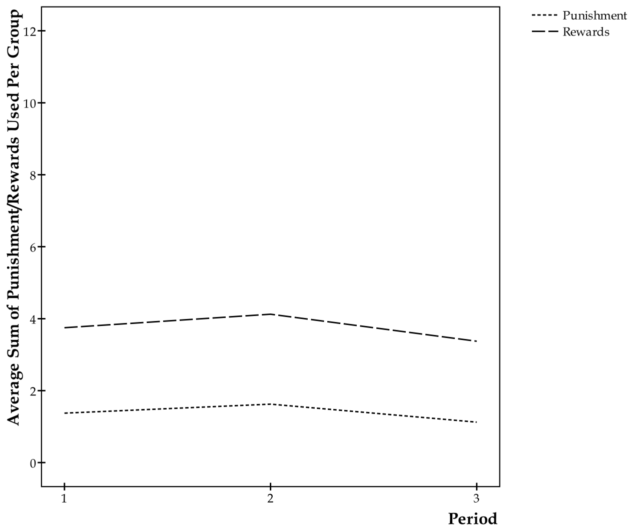


Figure 2 Average number of five minute punishment (reward) blocks sent per group in part 1 of the Punishment (Reward) treatments.

Two results are evident from the figure.

Result 4 Rewards are allocated more frequently than punishments.

Support for result 4: A Mann-Whitney test, taking the sum of five minute blocks used per group over the three periods as an independent observation ($N_1 = 8, N_2 = 8, p = 0.065$), rejects the hypothesis that the number of reward and punishment blocks assigned is equal (at $p < 0.10$). ■

Result 5 There is no trend in punishment or reward use over the three periods of part 1.

Support for result 5: Consider a Wilcoxon matched pairs test where each matched pair consists of a group's total number of punishment or reward blocks assigned in period 1 and in period 3. Each group is taken as one independent observation. In the Punishment (Reward) treatment, there is no significant difference in the number of five minute blocks assigned between period 1 and period 3, as the p-value of the relevant Wilcoxon matched pairs test (with $N_1 = N_2 = 8$) equals 0.593 (0.599). ■

On average, rewards are assigned more frequently than punishments, as is the case in laboratory experiments (see for example, Vyrastekova and van Soest (2008) and Rand et al. (2009)). The temporal pattern in the quantities of rewards assigned is also in line with laboratory results(e.g., Vyrastekova and van Soest (2008)), with reward assignments increasing over time. However, the same is not the case for punishment. In laboratory experiments, the number of punishments imposed tends to decrease over time (Fehr and Gächter (2000)), both when individuals are in groups of fixed membership or when they are grouped with new players in each period (in this case there can be no future benefit from punishing members of one's current group). However, here we find that punishments are used infrequently, even in the first period. Therefore, the lack of cooperation observed in our setting does not appear to be due to the relatively small number of periods in our experiment.

In laboratory environments, the use of punishments and rewards appear to be at least partly motivated by reciprocal considerations. Those who free-ride in the experiment's social dilemma stage are punished while those who cooperate are rewarded. We consider whether a similar pattern is observed in our study. The results for the use of punishment and reward options are reported in Tables 2 and 3, respectively. These two tables relate the quantity of punishment and reward sent by fishermen based on the number of fish that the sender and receiver have caught. For example, the top-left cell of Table 2 indicates that 3 five-minute punishment blocks, each block imposing a cost of 15 minutes to the recipient, were assigned by subjects who caught zero fish themselves to subjects who also caught zero fish. The number in square brackets (151) indicates the maximum possible quantity of punishment blocks that could have been assigned (if all subjects who caught zero assigned a quantity of one to all others that also caught zero).[3] The

[3]The maximum number of punishment or reward blocks that can be imposed in the experiment is equal to 8 groups of 4 group members who can use 3 blocks in 3 periods, which equals 288.

patterns in the table are summarized as result 6.

Result 6 Non-cooperators are more likely to receive punishment than cooperators, while cooperators are more likely to assign punishment. Cooperators are more likely to assign rewards than non-cooperators. The likelihood of receiving a reward is independent of cooperation level.

		Total quantity of punishment assigned classified by the sender's catch			
		0	1	2	Total
Recipient's catch of fish	0	3	0	3	6
		[151]	[44]	[16]	[211]
	1	20	1	0	21
		[43]	[8]	[5]	[56]
	2	6	0	0	6
		[16]	[5]	[0]	[21]
Total		29	1	3	
		[210]	[57]	[21]	[288]

Table 2 Punishment assignments, by catch of sanctioner and recipient. Numbers in brackets represent maximum possible number of assignments.

		Total quantity of reward assigned classified by the sender's catch			
		0	1	2	Total
Recipient's catch of fish	0	52	16	0	68
		[169]	[39]	[15]	[223]
	1	15	4	0	19
		[38]	[6]	[3]	[47]
	2	5	0	0	5
		[15]	[3]	[0]	[18]
		72	20	0	
Total		[222]	[48]	[18]	[288]

Table 3 Reward assignments, by catch of rewarder and recipient. Numbers in brackets represent maximum possible value.

Support for result 6: Table 2 indicates that those who catch the fewest fish are the most prone to imposing punishments. Players who catch no fish do so in 29 of 210, or 13.8%, of possible cases, while those who catch at least one fish punish in 4 out of 78 possible instances (5.1%). Punishments imposed by those who catch no fish tend to be directed at non-cooperators. In 26 of 59 possible instances (44.1%), players who catch no fish punish another group member who has caught at least one fish. Those who catch two fish are not punished more often than those who catch one fish.

Similarly, Table 3 indicates that those fishermen who catch relatively few fish are also more inclined to send rewards to others. Players who caught 0 fish assign rewards in 72 of 222, or in 32.4% of all instances, while Players who caught 1 fish do so in 20 of 48, or in 41.7% of the opportunities. Those who caught 2 fish never reward others. The probability that a player receives a reward is rather independent of how many fish he caught: 30.4%, 40.4%, and 27.8% for those who caught 0, 1, and 2 fish respectively. ■

We now consider whether receiving a punishment or a reward has any effect on subsequent cooperation. Table 4 contains the estimates of a regression in which the dependent variable is subject i's change in catch $x_{i,t+1} - x_{i,t}$ (columns (i) and (iii)), or change in effort $e_{i,t+1} - e_{i,t}$ (columns (ii) and (iv)) from period t to period $t+1$. The independent variables are the sum of punishment or rewards received by subject i, namely $\sum_{j \neq i}^{N} p_{ji}$. In case of catch, the variable $x_{i,t} - x_{-i,t}$ is included. This variable takes into account the effects of regression to the mean, independent of the punishment or reward instrument. Note that all subjects are given information on each group member's catch, but not on each group member's effort. Therefore, the regression to the mean variable is not included for the case of effort. The table serves as the basis for result 7.

Result 7 Recipients of reward or punishment do not cooperate more in the subsequent period.

Support for result 7: Table 4 shows that neither the receipt of punishment or reward causes subjects to catch fewer fish, or exert less effort in the

Dependent variable: $x_{i,t+1} - x_{i,t}$ ((i) and (iii)), $e_{i,t+1} - e_{i,t}$ ((ii) and (iv))				
	(i)	(ii)	(iii)	(iv)
	Punishment		Reward	
$\sum_{j \neq i}^{N} p_{ji}$	0.358^{***}	0.038	0.128^{**}	0.024
	(0.071)	(0.021)	(0.061)	(0.044)
$x_{i,t} - x_{-i,t}$	-0.67^{***}		-0.355^{***}	
	(0.157)		(0.077)	
Constant	-0.216^{*}	-0.064^{***}	-0.235^{**}	-0.051
	(0.108)	(0.013)	(0.1)	(0.051)
N	64	64	64	64
R^2	0.3797	0.0396	0.2685	0.0114

Table 4 OLS model to estimate the determinants of a change in catch (columns (i) and (iii), or a change in effort (columns (ii) and (iv))). Standard errors, clustered at the group level, are reported between parentheses.***: significant at the 1%-level, **: significant at the 5%-level.

next period. Moreover, the signs of the punishment and reward coefficients are opposite of what is expected; the more punishment or reward a subject receives, the more fish he catches in the next period. The positive effects of punishment and reward on cooperation that appears in many laboratory experiments are not found here. ■

4 Conclusion

In our experiment, we find no evidence that punishment and reward opportunities increase cooperation. All of the available measures indicate that fishermen try as hard as they can to catch as many fish as they can in part 1 of the sessions, regardless of whether or not punishments or rewards are available. Furthermore, receiving a punishment or a reward fails to increase the subsequent level of cooperation.

The ineffectiveness of punishment in creating cooperation contrasts sharply with most previous laboratory studies, but is consistent with the recent work of Janssen et al. (2010). Their study, while conducted in the laboratory, has the most similar context to ours of any study. This suggests that the difference between most laboratory studies and ours may lie in features of the games played, rather than in the fact that ours was conducted outside the laboratory. Our setting shares many features with Janssen et al. (2010), including (a) that individuals engage in a real time activity, (b) that they can update their cooperation tendency in real time in response to cues about how much others are cooperating, and (c) that the experiment is framed as harvesting of a resource. Any one, or a combination of these features, may make punishment and reward less effective.

Another aspect of our setting, which is not present in Janssen et al. (2010), is that greater effort only imperfectly translates into a catch. The intuition that this might reduce the effect of punishment finds support in recent work by Ambrus and Greiner (2009) and Grechenig et al. (2010). They show that, in a laboratory experiment with imperfect monitoring, receiving punishment induces players to contribute less. The fact that imperfect monitoring is not a feature of Janssen et al. (2010) suggests that is not likely to be the main reason that we fail to observe cooperation here.

Despite its ineffectiveness in promoting subsequent cooperation, punishment is typically assigned to those who catch the most fish from those who catch the least. The link to earlier cooperation is less clear in the reward treatment, where rewards, though assigned mainly by cooperators, are given independently of how many fish the recipient has caught.

We observe that punishment is applied in a setting when it has no effect on subsequent cooperation, and by individuals who do not respond to receiving punishment themselves. The fact that ineffective punishments are still being imposed at positive cost to the punisher, suggests that our subjects have an inherent taste for punishing non-cooperators (de Quervain et al. (2004), Singer et al. (2006)), and a desire to reciprocate to prior unkind actions (Gintis (2000), Fehr and Fischbacher (2003), Gintis et al. (2003), Boyd and Richerson (1992), Henrich and Boyd (2001), Sethi and Somanathan (1996) Gintis (2000)).

The results underscore that characteristics of the setting and the participants can influence the effectiveness of decentralized punishment in increasing cooperation. Prior research, for example Ostrom (1990) and Casari and Plott (2003), shows that decentralized punishment can promote cooperation in field settings. However, it is also clear that there exist many unresolved social dilemmas in the field. Our results suggest that for punishment or reward opportunities to be effective, interplay with other factors is necessary. In our setting, one problem might be that there is status associated with the activity that imposes the negative externality. Our fishermen might value skillful fishing, and even though fishing success reduces the payoff of others, these others may feel that the players who successfully catch fish are entitled to a higher payoff. In other words, some participants may not view the norm of cooperation as the norm that punishment should enforce. The norm that we intend to create, one of maximizing social surplus, may not override the prevailing norm at the fishing site, which is to try to catch as many fish as possible. Thus, some punished parties may interpret punishment for catching fish as inappropriate and respond by trying even harder to catch more fish. A similar problem may inhibit cooperation in other field settings, especially those in which there are very strong monetary rewards or social status from achieving the highest payoff among group members, such as some cases of the extraction of a valuable resource, price competition in a market, a patent race, or competing for a high course grade or a job promotion.

References

Ambrus, A. and Greiner, B. (2009). Imperfect public monitoring with costly punishment - an experimental study. *Working paper.*

Anderson, M. and Putterman, L. (2006). Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. *Games and Economic Behavior*, *54*, 1–24.

Boyd, R. and Richerson, P. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, *13*, 171–195.

Carpenter, J. (2007). The demand for punishment. *Journal of Economic Behavior and Organization*, *62*, 522–542.

Casari, M. and Plott, C. (2003). Decentralized management of common property resources: experiments with a centuries-old institution. *Journal of Economic Behavior and Organization*, *51*(2), 217–247.

Chaudhuri, A. (2010). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, *14*, 47–83.

de Quervain, D., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. *Science*, *27*(5688), 1254–1258.

Falk, A. and Heckman, J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, *326*, 535–538.

Fehr, E. and Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*, 785–791.

Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, *90*(4), 980–994.

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

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, *206*, 169–179.

Gintis, H., Bowles, S., Boyd, R., and Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, *24*, 153–172.

Grechenig, K., Nicklisch, A., and Thöni, C. (2010). Punishment despite reasonable doubt: A public goods experiment with sanctions under uncertainty. *Journal of Empirical Legal Studies*, *7*(4), 847–867.

Harrison, G. and List, J. (2004). Field experiments. *Journal of Economic Literature*, *42*(4), 1009–1055.

Henrich, J. and Boyd, R. (2001). Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, *208*, 79–89.

Janssen, M., Holahan, R., Lee, A., and Ostrom, E. (2010). Lab experiments for the study of social-ecological systems. *Science*, *328*, 613–316.

Ledyard, J. (1995). *Public Goods: A Survey of Experimental Research*. Princeton University Press.

Levitt, S. D. and List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *Volume 21*(2), 153–174.

Levitt, S. D. and List, J. A. (2008). Homo economicus evolves. *Science*, *319*(5865), 909–910.

Nikiforakis, N. and Normann, H. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics*, *11*, 358–369.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Actions*. Cambridge University Press.

Ostrom, E., Walker, J., and Gardner, R. (1992). Covenants with and without a sword: Self governance is possible. *American Political Science Review*, *86*(2), 404–417.

Rand, D., Dreber, A., Ellingsen, T., Fudenberg, D., and Nowak, M. (2009). Positive interactions promote public cooperation. *Science*, *325*, 1272–1275.

Sefton, M., Shupp, R., and Walker, J. (2007). The effects of rewards and sanctions in provision of public goods. *Economic Inquiry*, *45*(4), 671–690.

Sethi, R. and Somanathan, E. (1996). The evolution of social norms in common property resource use. *American Economic Review*, *86*, 766–788.

Singer, T., Seymour, B., O'Doherty, J., Stephan, K., Dolan, R., and Frith, C. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, *439*, 466–469.

Stoop, J., Noussair, C., and van Soest, D. (2010). From the lab to the field: Cooperation among fishermen. *Working paper*.

Vyrastekova, J. and van Soest, D. (2008). On the (in)effectiveness of rewards in sustaining cooperation. *Experimental Economics*, *11*, 53–65.

Yamagishi, T. (1986). The provision of a sanctioning system as a public good. *Journal of Personality and Social Psychology*, *51*(1), 110–116.

A Statistical analysis of effort on catch

In this appendix, we show the correlation between catch and effort, as measured by the number of times a fisherman casts his rod per minute. We find a positive correlation: More effort leads to more fish, and therefore we interpret our measure of effort as a legitimate proxy for cooperativeness. An ordered Probit model is used, taking a fisherman's catch and effort in each of the four periods as an independent observation.

As Table 5 shows, the coefficient of Effort is highly significant. The quadrant dummy variables correct for the spot at which a fisherman is fishing. Two of the quadrant variables are significant, indicating that the probability to catch fish at some spots is higher than at other spots.

Dependent variable: Number of fish caught in a period	
Effort	0.871***
	(0.158)
Quadrant Fixed Effects	Yes
N	448
pseudo-R^2	0.0411

Table 5 Relationship between individual effort and individual catch. Standard errors, clustered at the subject level, are reported between parentheses. ***: significant at the 1%-level.

B Questionnaire results

In this section we will present the questionnaire and the results that we used to determine whether a social dilemma exists.

B.1 The questionnaire

Dear Fisherman,

On behalf of Tilburg University we would like your cooperation to fill in a questionnaire. We ask you to indicate which of the following two options is your preferred option.

Option 1	Option 2
Fish 1.5 hours at maximally 6 fish	Wait for 1.5 hours at your fishing spot and then fish for 2.5 unlimitedly
Fish 1.5 hours at maximally 6 fish	Wait for 1.5 hours at your fishing spot and then fish for 2.5 unlimitedly, receiving €0.50 for each fish caught
Fish 1.5 hours at maximally 6 fish	Wait for 1.5 hours at your fishing spot and then fish for 2.5 unlimitedly, receiving €1 for each fish caught
... ...	... ...
Fish 1.5 hours at maximally 6 fish	Wait for 1.5 hours at your fishing spot and then fish for 2.5 unlimitedly, receiving €15 for each fish caught

B.2 Results of the questionnaire

In all of our treatments, the social optimum is induced only when fishermen prefer to wait for 1.5 hours in order to fish unlimitedly for 2.5 hours and receiving €2 per fish. The following table shows the results of the previous questionnaire. In total, we have surveyed 21 fishermen.

Willingness to accept to choose option 2	€0	€0.50	€1	€1.50	>€2
Number of Fishermen	14	3	1	1	2

Table 6 Results of the survey

As can be seen, 19 out of 21 fishermen claim to be better off when they wait for 1.5 hours and then fish unconstrained for 2.5 hours and be paid €2 for each fish caught. We are therefore confident that our experimental parametrization induces a social dilemma.