



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

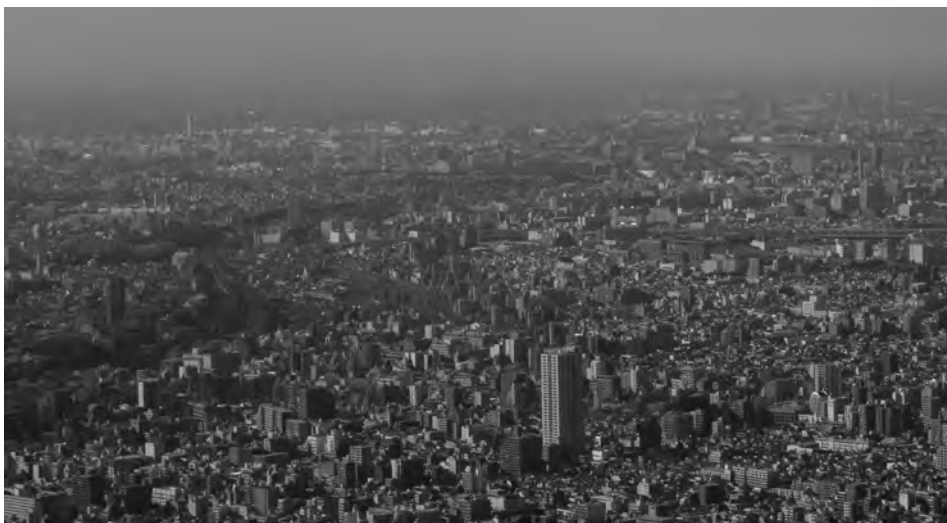


The handle <http://hdl.handle.net/1887/45536> holds various files of this Leiden University dissertation

Author: Molenmaker, Welmer E.

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

Issue Date: 2017-01-19





Chapter 2

On the willingness to costly reward cooperation and punish non-cooperation: The moderating role of type of social dilemma

This chapter is based on: Molenmaker, W. E., De Kwaadsteniet, E. W., & Van Dijk, E. (2014). On the willingness to costly reward cooperation and punish non-cooperation: The moderating role of type of social dilemma. *Organizational Behavior and Human Decision Processes*, 125(2), 175-183. doi: 10.1016/j.obhdp.2014.09.005

■ Abstract

Sanction opportunities are often introduced to promote cooperative choice behavior. Experimental studies have repeatedly demonstrated that the use of both rewards and punishments can indeed effectively increase cooperation. However, research has only recently begun to identify the determinants of the willingness to sanction. We investigate the use of costly sanctions to promote cooperation in the context of social dilemmas. We argue and demonstrate that people's willingness to costly reward and punish is not only determined by the type of sanction (reward versus punishment) but is also moderated by the type of social dilemma people face (public good dilemma versus common resource dilemma). In two experiments, we demonstrate that people punish less often and to a lesser extent than they reward, especially in a public good dilemma compared to a common resource dilemma.

■ Introduction

As a member of groups, organizations, and societies, we frequently encounter situations that require us to cooperate with others. This may involve cooperation with relatives, colleagues, and neighbors, but also with strangers. In many of these situations, we may be confronted with others who do not feel inclined to cooperate. The fact that groups often include members who do not cooperate can be detrimental to the collective. For example, group performance may suffer from group members who expect that others will compensate for their lack of effort (i.e., free-riders), organizations may be less efficient when employees work independently of each other, and the natural environment is jeopardized by the many environmental-unfriendly choices people make. Thus, the welfare of the collective is often influenced by the individual choices people make, either positively (in case of cooperative choice behavior) or negatively (in case of non-cooperative choice behavior).

From a collective point of view, it comes as no surprise that authorities often employ sanctions to promote cooperative choice behavior. Sanctions can either be positive means to increase cooperation (i.e., rewards, such as a bonus, price, or privilege) or negative means to decrease non-cooperation (i.e., punishment, such as a fine, penalty, or restriction). Research from a variety of disciplines, such as social psychology (e.g., Blau, 1964; Eisenberger, Lynch, Aselage, & Rohdieck, 2004; Gouldner, 1960; Komorita & Barth, 1985; Thibaut & Kelley, 1959; Wit & Wilke, 1990; Yamagishi, 1986, 1988), organizational behavior (Cropanzano & Mitchell, 2005), and economics (e.g., Abbink, Bolton, Sadrieh, & Tang, 2001; Brosig, Weimann, & Yang, 2004; Fehr & Gächter, 2000, 2002; Rand, Dreber, Ellingsen, Fudenberg, & Nowak, 2009) have repeatedly shown that both means can effectively promote cooperation (for an overview, see Balliet, Mulder, & Van Lange, 2011). However, to effectively promote cooperative choice behavior, decision makers in control of rewards and punishments should of course first be willing to provide and impose them. After all, sanction opportunities can only show their effect if they are actually administered.

In this chapter, we address this important aspect of implementing sanction opportunities by investigating people's willingness to costly reward cooperation and costly punish non-cooperation. Specifically, we focus on two factors that may determine whether people consider sanctioning the appropriate course of action (see March, 1994; Messick, 1999): the type of sanction they can administer and what kind of (non-)cooperative choice behavior they face.

The need for sanctions

To investigate the willingness to sanction, it is first important to understand why it is often necessary for authorities to promote cooperative choice behavior. Although cooperation is socially beneficial, the occurrence of mutual cooperation is not self-evident. After all, the collective interest does not necessarily coincide with the personal interest (Hardin, 1968; Olson, 1965; Samuelson, 1954). As a consequence, people often face the dilemma whether to further the collective interest or their personal self-interest. Situations that revolve around such a conflict are often referred to as social dilemmas (for overviews, see Parks, Joireman, & Van Lange, 2013; Van Lange, Joireman, Parks, & Van Dijk, 2013; Weber, Kopelman, &

Messick, 2004). Social dilemmas constitute the context in which we investigate the willingness to sanction.

Two important types of social dilemmas are the *public good dilemma* and the *common resource dilemma* (Camerer, 2003; Dawes, 1980). Public good dilemmas model the problem of realizing public goods from which all people may benefit, irrespective of whether or not they individually contributed to their provision. Blood transfusions, public broadcasting, and medical care are all real-world examples of public good dilemmas. For an individual it is more profitable not to contribute because contributing is costly, and eventually everybody can make use of public goods. However, if too many people choose not to contribute, public goods cannot be provided and the collective will be worse off than if people would decide to contribute. In common resource dilemmas, by contrast, people have to decide whether or not to restrict harvesting from scarce common resources. For example, energy conservation, overfishing, and water scarcity are all problems arising from excessive consumption. While it is in the individual's interest to consume from such common resources, these resources will deplete if people do not constrain their harvesting.

The use of sanctions is usually proposed as a means to promote cooperation (Hardin, 1968; Olson, 1965), and early social dilemma research on the willingness to sanction showed that there are also people willing to incur costs for punishments if they expect or fear that others will defect (Yamagishi, 1986, 1988). In fact, people prefer societies with sanctioning institutions over sanction-free societies (Güerker, Irlenbusch, & Rockenbach, 2006). Furthermore, the level of cooperation increases when there are people present who are prepared to sanction at their own expense (e.g., Fehr & Fischbacher, 2004; Fehr & Gächter, 2000, 2002; Milinski, Semmann, & Krambeck, 2002; Ostrom, Walker, & Gardner, 1992; Rand et al., 2009; Sefton, Shupp, & Walker, 2007; Walker & Halloran, 2004). Consequently, the willingness to costly reward cooperators and punish non-cooperators is considered to be a prerequisite for cooperation (e.g., Boyd & Richerson, 1992; Fehr & Rockenbach, 2004; Gintis, 2000; Gintis, Bowles, Boyd, & Fehr, 2003; Gintis, Henrich, Bowles, Boyd, & Fehr, 2008). Altogether, the general picture emerging from these earlier studies is that there are indeed people who are willing to provide and impose sanctions to promote cooperation, even if it is costly to do so.

Whereas people may use costly sanctions, very little research focused on the distinction between the willingness to use rewards for cooperation versus punishments for non-cooperation (for an exception, see Sutter, Haigner, & Kocher, 2010; see also Molm, 1997; Wang, Galinsky, & Murnighan, 2009). Are people equally willing to costly reward cooperation as they are willing to costly punish non-cooperation, or do they have a preference for one over the other? This question needs to be addressed to identify the determinants of people's willingness to administer sanctions in social dilemmas. In the present chapter, we propose that people's willingness to use punishments may differ markedly from their willingness to use rewards. More importantly, we argue and show that people's willingness to reward and punish depends on whether they face a public good dilemma or a common resource dilemma.

The willingness to costly reward and punish

The majority of research on costly sanctioning focused on punishment of non-cooperation, thereby neglecting the possibility to reward cooperation. This is surprising since rewarding cooperation also proved to be an effective means to promote cooperation (Balliet et al., 2011). We argue that people may have a general preference for administering rewards over punishments as a means to promote cooperation. Why do we think this is the case? Research on the *do-no-harm principle* showed that, even if the overall benefit outweighs the harm done, people are reluctant to inflict harm on others (Baron, 1993, 1995; Baron & Jurney, 1993; Baron & Ritov, 1994; Ritov & Baron, 1990; Spranca, Minsk, & Baron, 1991; see also Van Beest, Van Dijk, De Dreu, & Wilke, 2005). The same reasoning may apply to the use of rewards and punishments since both are beneficial in the sense that they can enhance cooperation. However, only the use of punishments – in contrast with the use of rewards – implies that one directly inflicts harm to another person. Based on this reasoning, we thus propose that people may be more reluctant to punish than to reward (cf. Abbink, Irlenbusch, & Renner, 2000; Offerman, 2002).

The do-no-harm principle has never been related to costly sanctioning in social dilemmas. Some earlier studies, however, provide indirect evidence for the above reasoning. For instance a study by Sutter et al. (2010) showed that people are more supportive of sanctioning institutions that administer rewards than sanctioning institutions that administer punishments. Furthermore, research on the use of secondary sanctions demonstrated that people who punished non-cooperators were punished themselves, whereas people who rewarded cooperators were rewarded themselves (Cinyabuguma, Page, & Putterman, 2006; Denant-Boemont, Masclet, & Noussair, 2007; Kiyonari & Barclay, 2008; Milinski et al., 2002; Nikiforakis, 2008; Rand et al., 2009). Such secondary sanctioning suggests that people evaluate punishments negatively and rewards positively. We believe that such differences may also be observed for first-order sanctioning. In fact, we argue that people may be less willing to costly punish non-cooperative choice behavior than to costly reward cooperative choice behavior. More importantly, however, the willingness to sanction may not only be determined by the type of sanction (reward versus punishment) but may also be moderated by the type of social dilemma people face (public good dilemma versus common resource dilemma).

Sanctioning in public good dilemmas versus common resource dilemmas

Both public good dilemmas and common resource dilemmas refer to the same conflict of interests (i.e., self-interest versus collective interest), and can be structured as each other's equivalents in terms of payoffs (Camerer, 2003; Dawes, 1980). When it concerns the willingness to costly sanction, however, we argue that public good dilemmas and common resource dilemmas should certainly not be treated similarly because they differ in the way in which the initial property is distributed (e.g., Camerer, 2003; Dawes, 1980; Van Dijk & Wilke, 1997, 2000). In public good dilemmas, people initially possess property themselves (private property), and decide whether or not they contribute this property to a public good. In common resource dilemmas, the property is initially located in a common resource (collective property), and

people decide whether or not they consume from this common resource. As a consequence, the property rights in public good dilemmas are considered private, whereas the property rights in common resource dilemmas are considered collective (Van Dijk & Wilke, 1997; see also Van Dijk, Wilke, & Wit, 2003).

Prior research showed that property rights (e.g., people's perception that money they decide on is their own) may lower people's willingness to allocate parts of their property to others (e.g., Cherry, 2001; Hoffman, McCabe, Shachat, & Smith, 1994; Muehlbacher & Kirchler, 2009; Oxoby & Spraggon, 2008). This indicates that people who consider themselves to have rights over a property feel more entitled to retain it. As such, the initial distribution of property may be an important factor. In a study on ultimatum bargaining, Leliveld, Van Dijk, and Van Beest (2012) showed that bargainers felt less entitled, and were less willing to allocate money to themselves, if they felt that the money was initially owned by their opponent than if the property was initially owned by themselves. We argue that the same reasoning may not only apply to the own versus other's property distinction, but also to the distinction between private and collective property. In other words, people may feel they are more entitled to private property than to collective property (i.e., the starting point in public good dilemmas versus common resource dilemmas, respectively). Moreover, we reason that this perception is not egocentric in the sense that it only affects how people perceive their own rights, but that it is more general, and that people also assign these rights to others. Thus, people may also feel that others are more entitled to their own private property than to the collective property.

Importantly, we argue that the dissimilarity in property rights across public good dilemmas and common resource dilemmas may lead to a difference in how people respond when they observe cooperation or non-cooperation. More specifically, since cooperation in a public good dilemma requires one to give up private property, people may consider this highly commendable (and thus more rewardable). After all, in the public good dilemma, people freely give up their property rights to further the collective interest. By contrast, cooperation in the context of a resource dilemma primarily means that one keeps collective what was collective. Although this may be seen as commendable, it may be less in need of an explicit reward because one had no (personal) rights to the collective resource to start with. However, when one does infringe on collective property in a common resource dilemma, people may consider such non-cooperation more objectionable (and thus more punishable) than not giving up private property in a public good dilemma. Thus, we argue that people may be less willing to costly punish non-cooperative choice behavior and more willing to costly reward cooperative choice behavior in public good dilemmas than in common resource dilemmas.

This proposition connects to research by Janoff-Bulman, Sheikh, and Hepp (2009) on the attribution of credit and blame across proscriptive and prescriptive morality (see also Goodwin & Darley, 2012; Janoff-Bulman & Carnes, 2013). In their research, they distinguish between morality that prescribes *what to do* (prescriptive morality) and morality that proscribes *what not to do* (proscriptive morality). Although both forms of morality can serve collective interests to the same extent, Janoff-Bulman et al. (2009) argue that they derive from distinct underlying systems. Prescriptive morality is activation-based and focusses on *doing what is good*

(for others). Proscriptive morality, on the other hand, is inhibition-based and focusses on *not doing what is bad* (for others). As they argued, prescriptive morality has a less mandatory and less strict nature than proscriptive morality. In agreement with this, they noted that participants indicated greater disapproval (moral blame) for people who hurt rather than did not help others (Janoff-Bulman et al., 2009, Study 5). Moreover, participants indicated greater approval (moral credit) for people who helped rather than did not hurt others (Studies 6 and 7).

Since cooperation requires people *to do what is good* for the collective in public good dilemmas (i.e., giving up private property) and *not to do what is bad* for the collective in common resource dilemmas (i.e., not infringing on collective property), we can draw a parallel between these types of social dilemmas and the two forms of morality that Janoff-Bulman et al. (2009) distinguish (prescriptive and proscriptive morality). The difference in moral (dis)approval between prescriptive morality and proscriptive morality is in line with our reasoning that (non-) cooperative choice behavior may be punished less and rewarded more in public good dilemmas than in common resource dilemmas.

■ Experiment 2.1

As a first test of our ideas, we used a *third party sanction paradigm* (see Fehr & Fischbacher, 2004) in which participants observed the choice behavior of two persons in a one-shot social dilemma task. This social dilemma context was either presented as a public good dilemma or a common resource dilemma. Subsequently, participants could either costly punish or costly reward one of the persons. Thus, participants themselves were not involved in the social dilemma (third party perspective), they only had the opportunity to sanction others in either a public good dilemma or a common resource dilemma. Based on our reasoning, we expected that participants (1) would punish a low cooperator less than that they would reward a high cooperator and (2) that they would punish a low cooperator less and reward a high cooperator more in a public good dilemma than in a common resource dilemma.

Method

Participants and design

Participants were 122 students at Leiden University (92 women and 30 men; $M_{\text{age}} = 19.67$ years, $SD_{\text{age}} = 3.48$).¹ A 2 (Social Dilemma Type: Public Good versus Common Resource) $\times$ 2 (Sanction Type: Reward versus Punishment) between-participants factorial design was used. Afterwards, two participants indicated on an open question about the purpose of the experiment that they did not understand their task correctly and that they made a mistake

¹ For each experiment, we aimed to recruit as many participants as possible within the given time available in the lab (approximately two weeks per experiment).

on our main dependent variable. Therefore, we excluded the data of these participants from our analyses.²

Procedure

We invited participants to the laboratory to participate in an experiment on “group decision making”. Upon arrival we seated them in separate cubicles, each containing a personal computer that was used to give instructions and to register their responses. A computer automated procedure assigned the participants randomly to one of the four conditions.

In the instructions we informed participants they had to perform a joint task with two fellow participants whose identities were unknown. In the joint task, one person (named person C) would observe the other two persons (named person A and B) performing a task. Subsequently, this person would have the opportunity to assign decrement coins (*punishment* conditions) or increment coins (*reward* conditions) to one of the two persons. Random selection would ostensibly determine which role each participant got. All participants then learned that they were assigned to the role of observer (person C). The instructions also explained to the participants that their choices made in the joint task determined how much extra money they could earn on top of their initial participation fee. After the study was conducted, a lottery would randomly select participants that would actually receive this extra money.

After this, the task that person A and B would perform (i.e., the social dilemma context) was explained. The type of social dilemma was manipulated in such a way that the payoff structures of the two social dilemmas were identical (see e.g., Dawes, 1980; Van Dijk & Wilke, 1995, 1997; Van Dijk et al., 2003). In the *public good* conditions person A and B would be endowed with 100 coins worth €0.10 each (in total €10, which is approximately US \$13). They could contribute these coins to a common pool or keep the coins for themselves. The contributed coins in the common pool would be multiplied by 1.5 and divided equally among the two persons. The kept coins would accrue totally to oneself. In the *common resource* conditions person A and B could harvest maximally 100 coins each from a common pool of 200 coins (also worth €0.10 each). The harvested coins would accrue totally to oneself and the coins left in the common pool would be multiplied by 1.5 and divided equally among the two persons. After participants read the rules of the social dilemma task, we posed five practice questions to ensure comprehension of the task. For example, we asked participants how many coins the persons possessed (*public good* conditions) or how many the persons could maximally harvest from the common pool (*common resource* conditions). After answering each question, the correct answer was disclosed.

Next, participants read instructions about their role in the joint task. In the *punishment* conditions participants learned that they would be endowed with 200 coins (also worth €0.10 each) which they could assign as decrement coins to one of the persons in the task (we never used the word ‘punishment’). The assigned decrement coins would be multiplied by

² Inclusion of the data of the excluded participants did not alter the pattern of results.

3 and subtracted from the individual outcome of the person concerned. The kept coins would accrue totally to oneself. The instructions in the *reward* conditions were identical, the only difference being that the participants learned that they could assign their coins as increment coins and that these assigned coins would be multiplied by 3 and added to the individual outcome of the person concerned (we also never used the word ‘reward’). After participants read the instructions, we posed seven practice questions to ensure comprehension of their role in the joint task. For example, we asked participants how many coins would be subtracted from (*added to*) the individual outcome of the person concerned when they would assign 1 decrement coin (*increment coin*). After answering each question, the correct answer was disclosed.

After the practice questions, the joint task started and participants had to wait until person A and B finished reading and made their decision. This took about 2 minutes. Next, participants received (bogus) feedback about the individual decisions of the two persons. In the *public good* conditions participants learned that person A contributed 80 coins to the common pool (i.e., high cooperator) and person B 20 coins (i.e., low cooperator). In the *common resource* conditions participants learned that person A harvested 20 coins from the common pool (i.e., high cooperator) and person B 80 coins (i.e., low cooperator). After receiving the feedback about the individual decisions of the two persons, participants first had to decide if and to whom they wanted to assign coins as decrement coins in the *punishment* conditions or as increment coins in the *reward* conditions. When participants decided to assign coins to person A or person B, they had to indicate how many coins they assigned. The number of coins participants could assign to a person ranged from 0 to 200 coins.

Finally, we checked the type of social dilemma and type of sanction manipulations. Besides, participants had to fill in questions to measure their comprehension and the believability of the joint task. At the end of the experiment, we thoroughly debriefed, paid (1 course credit or €3 monetary compensation) and thanked the participants. When the experiment was performed by all the participants, a lottery randomly selected five participants who received their actual earnings from the joint task.

Results

Manipulation checks

To check the manipulation of type of social dilemma, we asked participants whether the two persons (A and B) could harvest or contribute coins in the joint task. All participants except one (i.e., 99.2%) answered this question correctly.³ To check the manipulation of type of sanction, we asked participants whether they could assign decrement coins or increment coins. All participants (i.e., 100%) answered this question correctly. Altogether, these results indicated that our manipulations were successful and we included the data of all 120 participants in the analyses.

³ We included the data of these participants in our experiments because they did not indicate that they made a mistake on our main dependent variables and exclusion of the data did not alter the pattern of results.

Choice to sanction

As a first test we analyzed the influence of the Social Dilemma Type and the Sanction Type on the proportion of participants who chose to sanction ($N = 120$). A binary (Sanction Choice: 0 = not sanctioned, 1 = sanctioned) logistic regression yielded a significant Sanction Type main effect ($B = 1.22$, $SE = 0.47$, Wald ($df=1$) = 6.82, $p = .009$, Odds Ratio = 3.40, 95% CI [1.36, 8.53]) and a significant Social Dilemma Type $\times$ Sanction Type interaction effect ($B = 3.42$, $SE = 1.25$, Wald ($df=1$) = 7.49, $p = .006$, Odds Ratio = 30.63, CI [2.64, 355.22]). The Social Dilemma Type main effect was non-significant ($B = 0.15$, $SE = 0.45$, Wald ($df=1$) = 0.12, $p = .733$, Odds Ratio = 1.16, CI [0.49, 2.79]).

In line with our expectations, we can conclude that the proportion of participants choosing to punish (66.1%) was lower than the proportion of participants choosing to reward (86.8%). In addition, the results indicated that the proportion of participants choosing to punish was lower in the public good dilemma (53.3%) than in the common resource dilemma (79.3%), $\chi^2(1) = 4.44$, $p = .035$, Odds Ratio = 3.35, CI [1.06, 10.59]. In contrast, the proportion of participants choosing to reward was higher in the public good dilemma (96.8%) than in the common resource dilemma (76.7%), $\chi^2(1) = 5.41$, $p = .02$, Odds Ratio = 9.13, 95% CI [1.05, 79.54]. See Table 2.1 for the frequencies.

Table 2.1. Number of participants who sanctioned and not sanctioned, as a function of Sanction type and Social dilemma type (Experiment 2.1)

	Punishment		Reward	
	Public Good	Common Resource	Public Good	Common Resource
Sanctioned	16	23	30	23
Not sanctioned	14	6	1	7
Total	30	29	31	30

The choice whether or not participants wanted to sanction also involved deciding *who* they wanted to sanction. Not surprisingly, a Chi-Square test ($N = 92$) showed that the large majority of participants chose to sanction the high cooperator (92.5%) in the reward conditions and the low cooperator (94.9%) in the punishment conditions, $\chi^2(1) = 69.35$, $p < .001$, Odds Ratio = 226.63, 95% CI [39.37, 1304.48].⁴

Sanction size

A 2 (Public Good versus Common Resource) $\times$ 2 (Reward versus Punishment) ANOVA on the number of coins ($N = 120$) yielded a significant Sanction Type main effect

⁴This indicates that the majority of participants did not sanction anti-socially (Herrmann, Thöni, & Gächter, 2008; Parks & Stone, 2010).

($F(1,116) = 30.29, p < .001, \eta^2 = .19, 90\% \text{ CI } [.10, .29]$) and a significant Social Dilemma Type $\times$ Sanction Type interaction effect ($F(1,116) = 8.60, p = .004, \eta^2 = .06, \text{ CI } [.003, .15]$). The Social Dilemma Type main effect was non-significant ($F(1,116) = 1.763, p = .187, \eta^2 = .01, \text{ CI } [.00, .06]$). The Sanction Type main effect indicated that the size of the punishments ($M = 11.47, SD = 11.94$) was smaller than the size of the rewards ($M = 30.31, SD = 24.29$), which explained 19% of the variance.

The Social Dilemma Type $\times$ Sanction Type interaction effect explained 6% of the variance. To break down this interaction effect we performed simple-effect analyses. This revealed that the size of the rewards was significantly larger in the public good condition than in the common resource condition ($F(1,116) = 9.23, p = .003, \eta^2 = .06, 90\% \text{ CI } [.01, .14]$), while the size of the punishments in the public good condition and the common resource condition did not differ significantly ($F(1,116) = 1.27, p = .263, \eta^2 < .01, \text{ CI } [.00, .54]$). The simple-effect analyses also showed that in the public good condition the size of the punishments was significantly smaller than the size of the rewards ($F(1,116) = 36.18, p < .001, \eta^2 = .23, \text{ CI } [.13, .33]$). In the common resource condition the size of the punishments was smaller than the size of the rewards ($F(1,116) = 3.25, p = .074, \eta^2 = .02, \text{ CI } [.00, .08]$), although this difference was only marginally significant and the confidence intervals indicated that the precision of this estimation seemed low. See Table 2.2 for the mean number of coins and standard deviation per condition.

Table 2.2. *Number of coins assigned, as a function of Sanction type and Social dilemma type (Experiment 2.1)*

	Punishment		Reward		Overall	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Public Good	8.80	12.25	37.42	25.72	23.34	24.73
Common Resource	14.24	11.15	22.97	20.64	18.68	17.10
Overall	11.47	11.94	30.31	24.29	21.05	21.36

Discussion

In line with our reasoning, Experiment 2.1 demonstrated that participants were less willing to costly punish non-cooperative choice behavior than that they were willing to costly reward cooperative choice behavior. Furthermore, they were less willing to costly punish non-cooperative choice behavior and more willing to costly reward cooperative choice behavior in the public good dilemma than in the common resource dilemma. Thus, the results of Experiment 2.1 provide first evidence for our idea that people's willingness to costly reward and punish depends on the social dilemma context. This shows that people are less willing to punish those who do not cooperate by not giving up private property in public good dilemmas than those who do not cooperate by infringing on collective property in common resource dilemmas. Furthermore, people are more willing to reward those who cooperate by giving

up private property in public good dilemmas than those who cooperate by not infringing on collective property in common resource dilemmas.

Although the findings of Experiment 2.1 are in line with our reasoning, it should be noted that the difference in willingness to sanction between the public good dilemma and the common resource dilemma seemed to emerge particularly in the choice whether or not to sanction and the size of the rewards, and less so in the size of the punishments. Put differently, although fewer participants used punishments in the public good condition than in the common resource condition, the size of the punishments did not differ significantly between both conditions. This result is in line with earlier findings by Cubitt, Drouvelis, and Gächter (2011), who also found no difference in punishment *size* between both types of social dilemmas. Since people are reluctant to harm others (e.g., Baron, 1993, 1995; Baron & Jurney, 1993) and the participants' choice options were limited to punishment only in the punishment conditions (such as in the experiment by Cubitt et al., 2011), a possible explanation might be that the participants' willingness to punish a non-cooperator was so low, that a potential difference in the punishment size between both types of social dilemmas could not occur (i.e., a floor effect).

Nevertheless, our results – as well as many other studies (e.g., Fehr & Gächter, 2002; Yamagishi, 1986, 1988) – showed that there are indeed people who are willing to costly punish non-cooperation, even though the punishments they administered were small in size. At first glance this might seem to contradict the do-no-harm principle (e.g., Baron, 1995), which states that people are reluctant to inflict harm on others. However, the fact that participants' choice options were limited to one type of sanction only, might provide an explanation for these results. Note that under these conditions, a decision not to punish in Experiment 2.1 means that one does not sanction at all. Thus, choosing not to punish would mean that one would not respond to the observed inequality. It may very well be that – despite their reluctance to inflict harm on others – participants felt it would be undesirable if they would not react at all (for other research indicating that people are willing to reduce inequality in outcomes, see e.g., Bazerman, White, & Loewenstein, 1995; Fehr & Schmidt, 1999; Schroeder, Steel, Woodell, & Bembek, 2003; Tricomi, Rangel, Camerer, & O'Doherty, 2010). The solution would then be to punish, but punish mildly.

In our second experiment, we did present participants the opportunity to choose between rewarding and punishing. We anticipated that within an opportunity to address the observed inequality without having to punish, we might observe a greater reluctance to use the punishment option. Furthermore, this procedure increased the validity of our research since in real-life people are often able to choose between the use of punishments and rewards. Thus, in Experiment 2.2 we tested whether the type of social dilemma has an effect on the size of the sanctions when people can choose their own preferred type of sanction (reward or punishment). As such, we are able to measure people's default response to others' cooperative and non-cooperative choice behavior.

■ Experiment 2.2

In Experiment 2.2 we used the same third party sanction paradigm as in Experiment 2.1, with the only exception that participants now had the opportunity to choose between the use of costly rewards or costly punishments. Based on our reasoning, we again expected that participants (1) would punish a low cooperator less than that they would reward a high cooperator and (2) that they would punish a low cooperator less and reward a high cooperator more in a public good dilemma than in a common resource dilemma.

Method

Participants and design

Participants were 114 students at Leiden University (94 women and 20 men; $M_{\text{age}} = 18.99$ years, $SD_{\text{age}} = 1.73$) (see Footnote 1). We used a one-factor between-participants design with 2 levels (Social Dilemma Type: Public Good versus Common Resource).

Procedure

The procedure of this second experiment was almost identical to the procedure of Experiment 2.1. However, now participants could choose between assigning their coins as decrement coins (i.e., punishment) or as increment coins (i.e., reward). First participants indicated if and what kind of coins they wanted to assign.⁵ Only when participants decided to assign decrement/increment coins, they indicated to whom and how many of these coins they wanted to assign. Thus, after they chose the type of coin, they could only assign their coins as that type of coin to only one of the two persons and not, for example, as decrement coins to one person and as increment coins to the other person. As in Experiment 2.1, the number of coins participants could assign to a person ranged from 0 to 200 coins (also worth €0.10 each).

Results

Manipulation checks

To check the manipulation of type of social dilemma, we asked participants whether the two persons (A and B) could harvest or contribute coins in the joint task. All participants except two participants (i.e., 98.3%) answered this question correctly (see Footnote 3). This result indicated that our manipulation was successful and we included the data of all 114 participants in the analyses.

⁵ In the instructions and questions, the type of coins (decrement coins versus increment coins) were mentioned in a counterbalanced order. Including instruction order as variable in our analyses did not reveal any effect of the order of instructions on people's willingness to sanction and this did not influence the significance of other effects. Thus, significant effects remained significant and non-significant effects remained non-significant when instruction order was added as variable in the analyses.

Choice to sanction

First we analyzed the influence of the Social Dilemma Type on the proportion of participants who chose to sanction ($N = 114$). A binary (Sanction Choice: 0 = not sanctioned, 1 = sanctioned) logistic regression yielded a non-significant Social Dilemma Type main effect ($B = 0.27$, $SE = 0.79$, Wald ($df=1$) = 0.12, $p = .733$, Odds Ratio = 1.31, 95% CI [0.28, 6.13]). The proportion of participants who sanctioned (i.e. decided to either punish or reward) in the public good condition (i.e., 94.6%) was equal to the proportion of participants who sanctioned in the common resource condition (i.e., 93.1%). See Table 2.3 for the frequencies.

The choice whether or not participants wanted to sanction also involved deciding what type of sanction they wanted to use. We therefore compared the number of participants who punished and rewarded ($N = 107$), which showed that the proportion of participants choosing to punish (22.4%) was lower than the proportion of participants choosing to reward (77.6%), $Z = 13.70$, $p < .001$, Odds Ratio = 3.46, 95% CI [2.30, 7.70]. This lower proportion of participants choosing to punish compared to the proportion of participants choosing to reward was similar across the public good condition (punish = 19.9%; reward = 81.1%) and the common resource condition (punish = 26%; reward = 74%), $\chi^2(1) = 0.12$, $p = .73$, Odds Ratio = 1.51, CI [0.60, 3.77].

Furthermore, when participants chose to sanction, they subsequently indicated who they wanted to sanction ($N = 107$). As in Experiment 2.1, a Chi-Square test showed that the majority of participants chose to reward the high cooperators (92.5%) or punish the low cooperators (94.9%), $\chi^2(1) = 62.09$, $p < .001$, Odds Ratio = 103.12, 95% CI [20.39, 521.49] (see Footnote 4).

Table 2.3. *Number of participants who sanctioned (punished or rewarded) and not sanctioned, as a function of Social dilemma type (Experiment 2.2)*

	Public Good	Common Resource
Sanctioned	53	54
Punished	10	14
Rewarded	43	40
Not Sanctioned	3	4
Total	56	58

Sanction size

Since participants could choose what kind of coins they wanted to assign, the assigned coins could either be decrement coins or increment coins (or neither when participants choose not to assign any coins). Thus, participants chose a type of sanction in Experiment 2.2, whereas type of sanction was a between-participants factor in Experiment 2.1. To be able to compare

the sanction size of both types of sanction in one analysis, we coded decrement coins (i.e., punishment) as negative values (multiplied by -1), increment coins (i.e., reward) as positive values (multiplied by 1) and no coins assigned as zero. This enabled us to test our reasoning in one analysis since a negative mean implied that participants punished more than they rewarded, a positive mean implied that participants rewarded more than they punished, and a mean of zero implied that participants punished and rewarded equally.

A One-Way ANOVA on the coded number of coins ($N = 114$) yielded a significant Intercept ($F(1,112) = 47.58, p < .001, \eta^2 = .29, 90\% \text{ CI } [.18, .39]$). This indicated that the size of the rewards was larger than the size of the punishments since the value of mean coded number of coins differed positively from zero ($M = 24.98, SD = 39.42$), which explained 29% of the variance. In line with our expectations, a significant Social Dilemma Type main effect ($F(1,112) = 4.22, p = .042, \eta^2 = .03, \text{ CI } [.00, .09]$) was found. Thus, the results indicated that the preference to reward over punishment was stronger in the public good condition ($M = 32.59, SD = 46.98$) than in the common resource condition ($M = 17.64, SD = 28.97$). This explained 3% of the variance, although the confidence intervals indicated that the precision of this estimation is low. See Table 2.4 for the mean coded number of coins and standard deviation of the participants who chose to reward or punish per condition.

Table 2.4. *Number of coins assigned by participants who chose to sanction, as a function of Sanction type and Social dilemma type (Experiment 2.2)*

	Punishment			Reward		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Public Good	10	13.30	2.61	43	45.53	7.03
Common Resource	14	17.50	1.36	40	31.70	3.67
Overall	24	15.75	1.38	83	38.87	4.10

Discussion

The results of Experiment 2.2 were again in line with our reasoning. Thus, participants were less willing to costly punish non-cooperative choice behavior than that they were willing to costly reward cooperative choice behavior. Furthermore, they were less willing to costly punish non-cooperation than they were willing to costly reward cooperation in the public good dilemma than in the common resource dilemma. Experiment 2.2 thus further corroborated our idea that people's willingness to costly reward and punish depends on whether they face a public good dilemma or a common resource dilemma.

Experiment 2.2 showed strong support for our proposition that people have a general preference for administering rewards over punishments as a means to promote cooperation. When given the opportunity to choose between reward and punishment, the large majority of participants in both the public good dilemma and the common resource dilemma chose to

reward cooperation. As we suggested in the discussion of Experiment 2.1, having an option to choose reward rather than punishment as a way to address the observed inequality, can reduce the reliance on punishment because it allows one to address the observed inequality without inflicting harm. The fact that the proportion of participants who chose to punish in Experiment 2.2 (22.4%) was much lower than what we observed in the punishment condition of Experiment 2.1 (66.1%) accords with this notion. Note, however, that both patterns fit with our reasoning that people are reluctant to punish. Moreover, the observed difference in size of sanctions also corroborates the notion that people indeed consider choice behavior as more commendable (and thus more rewardable) and less objectionable (and thus less punishable) in public good dilemmas than in common resource dilemmas.

■ General discussion

The use of rewards and punishments is usually proposed as a means to promote cooperative choice behavior (Hardin, 1968; Olson, 1965). Whereas earlier research showed that people are often inclined to provide and impose sanctions (see e.g., Fehr & Gächter, 2002; Yamagishi, 1986; for an overview, see Balliet et al., 2011), no clear comparison has been made between the willingness to use costly rewards versus punishments. In the present chapter, we filled this gap in the literature by demonstrating that the willingness to sanction is not only determined by the type of sanction (reward versus punishment) but is also moderated by the type of social dilemma people face (public good dilemma versus common resource dilemma).

According to the do-no-harm principle, people are reluctant to inflict harm on others, even if the overall benefit outweighs the harm done (e.g., Baron, 1993, 1995; Baron & Jurney, 1993). In line with this principle, we observed that people punished non-cooperation less than they rewarded cooperation. In Experiment 2.1, we found that when people had to decide whether or not they wanted to punish, they sanctioned less often and to a lesser extent than people who had to decide whether or not they wanted to reward. Experiment 2.2 further showed that when people had the opportunity to choose between reward and punishment, the large majority chose to reward and subsequently rewarded to a greater extent than the people who chose to punish. Thus, when in the position to respond to others' choice behavior, people are reluctant to inflict harm on those who do not cooperate (i.e., punish), but they prefer to benefit those who do cooperate (i.e., reward).

Furthermore, we reveal a difference in costly sanctioning across public good dilemmas and common resource dilemmas. In two experiments, we demonstrated that people use punishments less often than rewards (Experiment 2.1) and to a lesser extent (Experiments 2.1 and 2.2) in public good dilemmas compared to common resource dilemmas. In public good dilemmas the property rights are private and in common resource dilemmas they are collective (Van Dijk & Wilke, 1997; see also Van Dijk et al., 2003). Against this background, people are less willing to punish those who do not cooperate by not giving up private property in public good dilemmas than those who do not cooperate by infringing on collective property in common resource dilemmas. Likewise, people are more willing to reward those who cooperate by giving up private property in public good dilemmas than those who cooperate by not

infringing on collective property in common resource dilemmas. Thus, these findings identify the allocation of property rights as an important determinant of people's willingness to reward cooperation and punish non-cooperation.

Implications, limitations, and directions for future research

Although one should always be cautious when generalizing experimental results to practice, two interesting implications may derive from our findings. First of all, the possibility of rewarding should not be overlooked by policymakers when implementing sanction opportunities in real-world social dilemmas. Earlier research indicated that costly rewarding comes with social benefits, while costly punishment has social costs (Kiyonari & Barclay, 2008; Milinski et al., 2002; Rand et al., 2009). In accordance with such secondary sanctioning research, our results (which pertained to first-order sanctioning) revealed that people have a relative preference for administering rewards over administering punishments. Although reward and punishment can both be effective means to enhance cooperation (Balliet et al., 2011), people will generally consider punishing the less appropriate course of action (see also March, 1994; Messick, 1999) and thus use punishments less often and to a lesser extent than rewards. Thus, when the opportunity to punish is not sufficiently used, implementing reward opportunities can be decisive to promote cooperative choice behavior.

The second important implication is that public good dilemmas and common resource dilemmas should not be treated as each other's equals, even in case of identical payoff structures. Although both social dilemmas appeal to the same conflict of interests (Camerer, 2003; Dawes, 1980), prior research already acknowledged that the type of social dilemma certainly has an impact on the people involved. For instance, the type of dilemma influences which social norm people adhere to (Van Dijk & Wilke, 1995), whether they experience social responsibility (Van Dijk & Wilke, 1997), and whether they prefer installing a leader (Rutte & Wilke, 1984; Rutte, Wilke, & Messick, 1987; Van Dijk et al., 2003). Our results extend such findings and emphasize the importance of taking the type of social dilemma (public good dilemma versus common resource dilemma) into consideration when implementing reward or punishment opportunities in real-world social dilemmas. Although people will generally prefer to reward cooperative choice behavior over punishment of non-cooperative choice behavior, this preference will be stronger in public good dilemmas than in common resource dilemmas.

Furthermore, the implications are also relevant for the research on costly sanctioning. For example, the majority of social dilemma research on costly sanctioning only focused on punishment of non-cooperation and was conducted within public good dilemmas (see Balliet et al., 2011). Our results reveal that the type of sanction (reward versus punishment) and the type of social dilemma (public good dilemmas versus common resource dilemmas) both determine people's willingness to sanction. This indicates that one should take both determinants into account when investigating the underlying considerations of costly sanctioning.

At this point, it is relevant to address some limitations of our research, and give some suggestions for future research. In our two studies, we used a third party sanction paradigm in which participants were not dependent on the choice behavior they observed (see Fehr &

Fischbacher, 2004). With this procedure, we ensured that the participants' interpretation of others' choice behavior was not colored by self-interest. At the same time, this procedure may constitute a limitation since one may wonder whether our idea that people are reluctant to use punishments would also be observed in situations in which people are part of the group themselves. People's willingness to punish non-cooperation may be higher when they are personally involved in the social dilemma because in such situations, revenge-like motives might justify the infliction of harm (see e.g., De Quervain et al., 2004). As such, it might be interesting for future research to test whether our finding that people are reluctant to administer punishments can be replicated using a second party sanction paradigm (see Chapter 3).

Another characteristic of our experimental paradigm is that we designed in such a way that we were able to compare people's willingness to administer rewards and punishments directly. Since the cost–consequence-ratio (i.e., 1:3) was identical for rewards and punishments, we merely varied the type of sanction. Thus, it was equally costly for participants to influence a person's outcome either positively (i.e., reward) or negatively (i.e., punishment). Nevertheless, an intrinsic element of rewards and punishments is the fact that the joint outcomes of a group rise with the use of rewards, whereas they drop with the use of punishments. As a result, people's preference to administer rewards over punishments may also reflect a preference to further the collective welfare (Sutter et al., 2010; see also Charness & Rabin, 2002). Such an efficiency perspective can, however, not explain why the preference to administer rewards over punishments is moderated by the type of social dilemma people face. It would, however, be a good idea for future research to further explore the differences in the underlying motives of rewarding versus punishing.

Finally, future research might want to explore to what extent discretion over what type of sanctions people can use may play a role in the use of sanctions. The results of our two studies suggest that while the majority of people generally opt for rewards instead of punishments, a substantial proportion of people may also administer punishments when this is the only type of sanction at their disposal. As we mentioned in the discussion of Experiment 2.1, participants in the punishment conditions may have punished because they felt inclined to reduce the inequality they observed (see e.g., Tricomi et al., 2010), even though this meant they had to inflict harm (i.e., punish). The opportunity to choose between reward and punishment, however, offers people an opportunity to respond to the inequality and at the same time refrain from punishment. It may therefore be interesting for future research to investigate people's willingness to sanction when the inequality is less obvious and they only have discretion over one type of sanction (see Chapters 3 and 4).

Conclusions

The present chapter addresses the importance of distinguishing between the willingness to costly reward and punish. As we were able to demonstrate, people's willingness to sanction is determined by the type of sanction (reward versus punishment) and the type of social dilemma they face (public good dilemmas versus common resource dilemmas). As a result, we hope that the present chapter represents an important step in identifying the determinants

of people's willingness to promote cooperative choice behavior with the use of sanctions. When one knows the determinants of costly sanctioning, one is able to predict how sanction opportunities will be used to promote cooperative choice behavior.