How to Organize Monitoring and Punishment: Experimental Evidence

Timo Goeschl

Beatrix Haberl

Alice Soldà

How to Organize Monitoring and Punishment: Experimental Evidence[*]

Timo Goeschl[1,2], Beatrix Haberl[3], and Alice Soldà[4]

[1]Alfred Weber Institute for Economics, Heidelberg University
[2]ZEW - Leibniz Institute of European Economic Research, Mannheim
[3]Department of Economics, University of Mannheim
[4]Department of Economics, Ghent University

September 20, 2023

Abstract

Punishment institutions for curtailing free-riding in social dilemmas rely on information about individuals' behavior collected through monitoring. We contribute to the experimental study of cooperation-enhancing institutions by examining how cooperation and efficiency in a social dilemma change in response to varying how monitoring and punishment are jointly organized. Specifically, we evaluate - against a no-monitoring baseline - combinations of two imperfect monitoring regimes (centralized vs. decentralized) and three punishment regimes (self- vs. peer- vs. delegated punishment) in a repeated public goods game. As hypothesized, we find that delegated punishment outperforms other punishment regimes, irrespective of the monitoring regime, both in terms of cooperation and efficiency. Monitoring, both centralized and decentralized, cannot raise cooperation relative to the baseline unless accompanied by a credible punishment. When combined with a punishment institution, both monitoring regime outperforms the baseline. (141 words)

JEL code: C72, C91, C92, D7, H41

Keywords: Compliance, monitoring, punishment, experiment

1   Introduction

Free-riding is a major problem in social dilemmas and, when unchecked, can be the source of significant welfare losses. An extensive experimental literature in economics has investigated institutions, mechanisms, and processes to limit free-riding. A key result of this literature has been that the presence of a punishment institution can significantly curtail free-riding and increase cooperation. When group members, for instance, have the option to monitor others in the group and punish them, such a peer punishment option increases group efficiency substantially and in a sustained way (Fehr and Gächter, 2000, 2002). Likewise, the threat of being punished by a monitoring centralized authority has been shown to be effective in preventing large-scale free-riding (*e.g.,* Falkinger et al., 2000; Tyran and Feld, 2006; Guillen et al., 2007; O'Gorman et al., 2009; Baldassarri and Grossman, 2011; Andreoni and Gee, 2012). How effective such punishment options are in curtailing free-riding has been shown to depend on a variety of factors that define the complex reality of a specific social dilemma.[1] One critical factor appears to be the organization of the punishment option (see Dannenberg and Gallier (2020) for a recent exhaustive review), in particular whether it is better for the punishment institution to rely on a decentralized design that requires peer action or on centralized design that delegates punishment action to another party, such as the police (*e.g.,* Sutter et al., 2010; Andreoni and Gee, 2012; Nicklisch et al., 2016; Gross et al., 2016).

The present paper contributes to this literature by studying - as its main novelty - the effect of *jointly* varying how the monitoring and the punishment stage of a social dilemma are organized. Like other researchers (*e.g.,* Page et al., 2005; Nicklisch et al., 2021), we are motivated by a rich empirical literature that has documented that groups come to very different arrangements on how to collect information about members' behavior and how to sanction free-riders (Ostrom et al., 1992). Our paper, in particular, explores centralization, *i.e.,* delegation to a single institution, and decentralization, *i.e.,* peer provision, both at the monitoring and the punishment stages of a social dilemma.

---

[1]Examples include the specific punishment costs (Egas and Riedl, 2008), the presence of a counter-punishment option (Nikiforakis, 2008; Nikiforakis and Engelmann, 2011), and the accuracy of monitoring to detect free-riding behavior (Grechenig et al., 2010; Ambrus and Greiner, 2012; Fischer et al., 2016).

By doing so, it integrates the more established literature on how to organize the punishment stage with a more recent literature highlighting the requisite role of the monitoring stage. There, it has been noted that variations in how information about group members' choices can be acquired and transmitted to other group members, *e.g.,* with respect to noise, delays, and costs, matters significantly (Ambrus and Greiner, 2012; Fudenberg and Pathak, 2010; Goeschl and Jarke, 2017). Closest to our paper are DeAngelo and Gee (2020) who study experimentally whether group or peer (nearest right-hand neighbor) monitoring will promote efficiency in the repeated public goods game with costly peer punishment. In particular, they are interested in whether the outcome of the comparison depends on monitoring, which is always costly and imperfect, being exogenously imposed or endogenously chosen. They find that peer and group monitoring are equally effective when exogenously imposed, but when the monitoring institution requires (near) unanimity to be implemented endogenously, peer monitoring performs better, in fact nearly as well as under exogenous peer monitoring.

Our design randomly assigns participants to one of seven conditions, six treatments and one baseline. In all conditions, participants play five rounds of a linear public goods game (PGG) in fixed groups of four. The Contribution Stage of the PGG is followed by an imperfect Monitoring Stage, which determines whether information about individual behavior in the contribution stage is made available to others and, if so, to whom. At the subsequent Punishment Stage, insufficiently cooperative ('non-compliant') group members can be punished by those who have learned about their behavior. Others learning about a group member's contributions requires that costly monitoring effort was expended on that member, and that this effort was successful. While this structure and a number of other choices such as ruling out anti-social or wrongful punishment closely resemble those in DeAngelo and Gee (2020), our design differs from theirs in several ways. The most important distinction is that we vary the punishment mechanism in combination to the monitoring regime. Another important distinction is that while their decentralized monitoring regime operates in a circular network (*i.e.,* group members can only monitor the group member to their right), our decentralized monitoring is such that every group

2

member can monitor all other group members.

In the baseline condition, group members can only know their own contributions and can only punish themselves. In the other six conditions, one treatment dimension varies the Monitoring Stage. There, monitoring can be either decentralized, *i.e.,* carried out individually by group members, or centralized, *i.e.,* delegated to a single mechanistic institution. Think of communal forest users (Rustagi et al., 2010) as an example of the first and CCTV monitoring by police (Piza et al., 2019) as an example of the second approach. The other dimension varies how the Punishment Stage is organized. In one variation, punishment is decentralized, *i.e.,* carried out individually by other group members (peers), as documented for pastoral communities in the Italian Alps (Casari and Plott, 2003; Casari, 2007). In another, it is centralized, *i.e.,* delegated to a single mechanistic institution, such as the police.[2] A third variation, included for completeness, is enforcement restricted to self-punishment.[3] These seven conditions each present different ways of organizing the monitoring of actions and the punishment of free-riding. For example, other traffic participants can observe someone else running a red light and report the incident, but only police can punish the offender. Conversely, international agreements, such as the Paris Agreement on Climate Change, sometimes delegate the monitoring of whether ratifying countries adhere to their obligations to dedicated secretariats, but sanctions need to be imposed by members individually. One of the seven conditions, combining centralized monitoring and decentralized punishment, also matches closely one of the four treatments implemented by DeAngelo and Gee (2020).

Our specific parameterization of the design ensures that none of the institutional choices can successfully enforce full compliance. Against this background, we have two sets of predictions for our design. Game-theoretic reasoning based on selfish and rational economic agents leads to predictions that effectiveness (contributions) and efficiency (payoffs net of institutional cost) will not differ across treatments. Behavioral reasoning based on existing experimental evidence leads to predictions that delegated punishment and centralized monitoring each outperform their decentralized alternative in terms of

---

[2]Other terms used for this is 'pool punishment' (Traulsen et al., 2012).

[3]While also skeptical about its potential, DeAngelo and Gee (2020) explicitly consider this possibility.

effectiveness and efficiency. The purpose of the experiment is to provide evidence to test these predictions.

Taking the design to 932 participants in an online environment, we find that, in line with our behavioral predictions, delegating punishment to an automatic mechanism performs best among the institutional arrangements tested: Contributions and net payoffs are strictly higher in treatments in which punishment is delegated compared to treatments in which punishment is decentralized, irrespective of whether monitoring is centralized or decentralized, or compared to self-punishment. Delegated punishment therefore constitutes our first order effect in terms of magnitude and significance. We also find that decentralized punishment, requiring costly monitoring, does better than the baseline in terms of contributions, but not in terms of net payoffs, in contradiction to our hypotheses.

Contrary to our expectations, we do not find evidence that centralization performs systematically better than decentralization in our experiment when it comes to monitoring. In fact, decentralized monitoring leads to higher net payoffs than its centralized alternative. This can be explained by the fact that centralized monitoring leads to higher monitoring expenses than decentralized monitoring (as every group member has to pay the monitoring fee in every round), thus harming net payoffs. For policy-makers focused on effective policies, these results can be read as a clear endorsement of delegated enforcement of a punishment option in social dilemmas, with the choice of the monitoring arrangements a secondary consideration. For the economist focused on efficiency, they can be read as providing considerable degrees of freedom in designing the institutional structure of monitoring and punishment that curtails free-riding, as long as some combination of decentralized and centralized components is in place.

Our discussion of these results acknowledges – alongside the important commonalities – the considerable diversity in the specifics under which experimental participants take their choices in the published literature. For example, in our design, participants become punishable only if they contributed less than 80 percent of their per-round endowment in the last round (so-called 'non-compliers') and were successfully monitored. This punishment eligibility differs from the 'lowest actually monitored contributor' policy

used elsewhere (*e.g.,* DeAngelo and Gee, 2020; Andreoni and Gee, 2012) in a number of important ways. Our results, and those of other papers, therefore need to be considered against the background of such specifics, as we explain in detail.

The paper proceeds as follows. In section 2, we describe our design and the experimental procedures. Section 3 develops the relevant theoretical considerations that give rise to a set of testable hypotheses. The empirical evidence and the results of testing these hypotheses are presented in section 4. Section 5 discusses our results in the context of the literature and concludes.

2   Experimental Design

Our experimental design[4] relies, at its core, on the repeated linear public good game in which we systematically vary (i) the monitoring regime (no monitoring vs. decentralized monitoring vs. centralized monitoring) and (ii) the punishment mechanism (self-punishment vs. peer-punishment vs. delegated punishment). These systematic variations result in a baseline with no monitoring and six treatments, displayed in Table 1. We use a between-subject design so that each experimental subject participates in one and only one of the treatments. In all treatments, participants interact in groups of four over five rounds, and the group composition remains fixed over the five rounds. Each round consists of four stages: A Contribution Stage, a Monitoring Stage, a Punishment Stage, and a Feedback Stage.

Table 1: Baseline and Treatments.

| | Monitoring regime | | |
| | No Monitoring | Decentralized Monitoring | Centralized Monitoring |
Punishment mechanism			
Self-punishment	SP-NM (baseline)	SP-DM	SP-CM
Peer punishment		PP-DM	PP-CM
Delegated punishment		DP-DM	DP-CM

[4]The complete instructions (including screenshots) are provided in Appendix B.

2.1 Contribution Stage

In the Contribution Stage, each group member receives an endowment of 10 tokens placed in a private Active Account and a deposit of 6 tokens placed in a private Passive Account. Group members then simultaneously decide how many tokens, between zero and ten, to contribute from their Active Account to a shared Group Account (*i.e.,* the public good). Each token contributed to the Group Account is multiplied by 1.6 and the product is divided equally among the four group members. Thus, for every token contributed to the Group Account, all four participants in the group receive 0.4 tokens in their Active Account, regardless of who contributed the token. In other words, the marginal per capita return (MPCR) from the public good is 0.4. Any token that a group member does not contribute to the Group Account remains in their Active Account.

Before making their contribution decision, participants are told that their contribution will be considered "high" if they contributed 8 tokens or more to the Group Account, and "low" if they contributed 7 tokens or less. This feature borrows from designs by Cherry and McEvoy (2013), Cherry and McEvoy (2017), and Barron and Nurminen (2020), which also set an 80 percent threshold that labels contributions at or above as "compliant" or "good" and vice versa. The reasoning is that at 80 percent of endowment, the threshold is higher than the typical average contribution level in PG experiments of around 40 to 60 percent (Chaudhuri, 2011), encouraging coordination at levels above those typically observed. At the same time, the threshold is below the social optimum of contributing 100 percent of the endowment. As a result, less ambitious changes in contribution behavior are required from the participants, making coordination at the threshold more plausible to participants (Barron and Nurminen, 2020). However, the feature of an absolute compliance standard is a departure from, for instance, Andreoni and Gee (2012) or DeAngelo and Gee (2020) who choose a relative standard. These designs follow Yamagishi (1986) that declare the lowest contributor among the group members as "non-compliant". Both approaches have real-world counterparts, and we are attracted to absolute standards on account of the many settings in which laws and regulations set standards exogenously and independently of others' behavior.

Formally, participant i's payoff in the Contribution Stage in round t is given by:

$$\pi_i^t = 10 - c_i + 0.4 \sum_{j=1}^{4} c_j$$

where c_i ($0 \leq c_i \leq 10$) is participant i's contribution to the Group Account and $\sum_{j=1}^{4} c_j$ the sum of contributions to the group account by all the group members, including participant i.

2.2 Monitoring Stage

In the Monitoring Stage, participants find themselves assigned to a baseline without monitoring or to one of two monitoring regimes. In the baseline, there is no monitoring (NM) and participants only receive information about (1) their own contribution to the Group Account, (2) whether or not their contribution was compliant, and (3) their group's total contribution to the Group Account. Participants receive this baseline information cost-free in all seven treatments.

In the two treatments with monitoring regimes, participants can become eligible for learning the contribution decisions and compliance status of other group members. Group members each receive one additional token in the Monitoring Stage. In the decentralized regime (DM), eligibility is elective and private: Participants decide individually whether to expend that token in order to become eligible for learning. Those who keep the additional token are not eligible and receive only the baseline information. Group members are not informed about others' decision regarding eligibility. In the centralized regime (CM), eligibility is mandatory: Participants have to expend that token and all become eligible for learning. This is akin to a tax-funded monitoring regime.

Being eligible for learning does not guarantee learning because in this experiment, like in DeAngelo and Gee (2020), monitoring is designed as imperfect. This means that efforts to monitor a group member succeed only two thirds of the time. Specifically, at the beginning of each round, each group member has a probability of $\frac{2}{3}$ to be monitored. If monitoring in that round is successful for a particular group member, his contribution

decision and compliance status will be disclosed to all eligible group members before the Punishment Stage. If monitoring in that round is unsuccessful for this group member, his contribution decision and compliance status will not be disclosed to any group member, eligible or not, before the Punishment Stage. This procedure has two implications: One is that whether monitoring of a group member in this round will be successful is determined at the group level. This means that all members eligible for learning in that round privately learn the same information. The second is that this information can contain the contribution decisions and compliance status of zero, one, two, three or all group members, depending on the outcome of the random draw. It is therefore possible that even eligible group members fail to learn anything during the Monitoring Stage, despite having expended their additional token on monitoring.

2.3   Punishment Stage

The design of the Punishment Stage is based on a deposit-refund mechanism. In this, we follow Cherry and McEvoy (2013, 2017), who experimentally implement the mechanism proposed in Gerber and Wichardt (2009). The mechanism makes use of the private Passive Account in which each participant has received a deposit of 6 tokens at the beginning of the Contribution Stage. This deposit is the target of possible punishment actions at the Punishment Stage.

To be punishable, a participant's behavior in the Contribution Stage of that round must have been successfully monitored and found to be non-compliant. Punishment, which is costless for the punisher, takes the form of deductions from the Passive Account. Through their punishment regimes, treatment conditions differ in who can punish punishable group members. There are three different punishment regimes: Self-punishment (SP), peer punishment (PP), and delegated punishment (DP). In SP treatments, non-compliant participants individually decide whether to forfeit the deposit in their Passive Account. If a non-compliant participant decides to forfeit their own deposit, all 6 tokens are deducted from their Passive Account. Compliant participants cannot forfeit their own deposit.

In PP treatments, the (possibly empty) set of potential punishers in a group consists of those eligible for learning. They decide, for each punishable group member, whether to deduct 2 tokens from their Passive Account. Those who are punishable can therefore lose up to 6 tokens from the deposit in their Passive Account in that round, namely if all three other group members are eligible to punish and do so. In DP treatments, all punishable group members automatically lose 6 tokens from their Passive Account if there is at least one group member expended their additional token on monitoring.

The design of the three punishment regimes, which restricts punishability to non-compliant group members who have been successfully monitored, precludes both anti-social and wrongful punishment. Compliant group members are safe from being targeted by punishers and from being mistaken for non-compliant group members.

2.4 Feedback Stage

In all treatments, each round concludes with a Feedback Stage. There, participants are informed about how many tokens (out of 6) are left on each group member's Passive Account. In addition, participants are informed about their total earnings in that round. These earnings are the payoff from Contribution Stage, any balance left over from the Monitoring Stage (at most 1 token), and the refund (balance) from the Passive Account from the Punishment Stage (at most 6 tokens).

2.5 Procedure

The experiment was programmed using oTree (Chen et al., 2016) and implemented online. We pre-registered to recruit 1,400 participants.[5] We recruited a total of 1,692 participants via the online platform prolific, of whom 1,452 completed the experiment. The group composition was kept constant over the five rounds of play (*i.e.,* partner design) and group members were identifiable by fixed IDs. Participants were randomly allocated to one of the seven treatment conditions. After reading through the instructions, participants had to answer five comprehension questions. Participants could only start the first round

---

[5]The preregistration can be found at https://osf.io/9frxh.

after answering all the comprehension questions correctly. Participants were paid for a randomly selected round with a £0.10 per 1 token conversion rate. Participants earned an average payoff of £6.99 (s.d. = £0.25), including a participation payment of £5.

3 Theoretical Predictions and Hypotheses

3.1 Theoretical predictions

Given the parametrization of the experiment, none of the institutional settings in the seven treatments can successfully deter free-riding by enforcing full compliance on selfish and rational economic agents. The three treatment conditions that contain a self-punishment (SP) regime rely, as a deterrence mechanism, on potentially non-compliant participants expecting at the Contribution Stage that they themselves will forfeit their own deposit at the Punishment Stage. Since doing so negatively affects their own payoffs, self-punishment will not occur. Consequently, selfish and rational group members will play the Nash equilibrium of zero contributions to the Group Account, regardless of the monitoring regime.

The two treatment conditions that contain a peer-monitoring regime (PP-PM, DP-PM) require that group members pay a 1-token fee in order to become eligible for learning about others' contributions and compliance status during the Monitoring Stage. This is a requisite for being able to punish non-compliant group members. Selfish and rational group members will not pay the fee. To see why, it suffices to understand that paying the fee, even if observable, cannot deter another group member's free-riding. Given others' contributions, being compliant reduces Contribution Stage payoffs by $((1-0.4)\times 8 = 4.8)$ tokens and would incur an expected punishment of no more than $(2/3\times 6 = 4)$ tokens, even if all other group members were able to jointly commit to paying the fee and punishment was delegated. Monitoring is therefore unproductive for raising own payoffs by enforcing other group members' compliance. No selfish and rational group member will therefore expend a token on becoming eligible to learn. Hence, participants' contributions to the Group Account will remain zero under peer monitoring, regardless of the enforcement

mechanism.

In the two treatment conditions that contain a centralized monitoring regime (PP-CM, DP-CM), monitoring occurs automatically as participants are forced to pay the access fee. Since reducing the deposits of group members who are found to be non-compliant does not impose additional costs on participants, peer enforcement can occur in equilibrium because rational participants are indifferent between reducing non-compliant group members' deposits or not. This will not result in compliance of selfish and rational group members, however, by the same argument as for PM. The benefits of non-compliance (4.8 tokens) exceed the (maximum) expected cost of non-compliance both under peer-enforcement and third-party enforcement (4 tokens). Participants will therefore contribute zero tokens to the Group Account under centralized monitoring as well, regardless of the punishment mechanism. This leads to the game-theoretic prediction that contributions to the Group Account are zero in all treatments for selfish and rational group members.

3.2 Behavioral hypotheses

Ample experimental evidence shows that institutional designs that are theoretically unable for deterring free-riding and non-compliance in social dilemmas can deliver high rates of cooperation and compliance in the laboratory (Chaudhuri, 2011; Dannenberg and Gallier, 2020). This is also true for social dilemmas with a deposit-refund mechanisms: Cherry and McEvoy (2017) find substantial compliance rates and contribution levels even when monetary deposits are small enough to be materially non-deterrent vis-à-vis free-riding. To take this evidence into account, we enrich the standard game-theoretic reasoning by empirical regularities established in the experimental literature to derive behaviorally augmented hypotheses, first, about the relative levels of cooperation to be expected in our treatments and, second, about the relative levels of efficiency (including the costs of the institution).

3.2.1 Contributions

Our first hypothesis about how contribution levels respond to joint variations in monitoring and punishment establishes a basic prediction about how a baseline with no monitoring and self-punishment will perform relative to six treatment conditions in which some form of monitoring or punishment is institutionally present. Specifically, we posit that compared to a SP-NM baseline, the other six treatments will not give rise to lower cooperation rates. The reasons is that in the SP-NM condition, group members cannot learn about others' contribution behavior at the individual level and have no reason to attach a positive probability to punishment for non-compliance. By contrast, the other six conditions, by making other group members' cooperation behavior at least potentially observable and by allowing for the possibility of punishment, allow at a minimum for behavioral mechanisms such as conditional cooperation to play out. This leads to our first hypothesis.

Hypothesis 1 *(Baseline vs. treatments)*: None of the six treatments leads to lower contributions than a baseline without monitoring.[6]

Our second hypothesis considers predictions on the effect of peer punishment and delegated punishment relative to a self-punishment option. While the extensive experimental literature consistently shows that both peer and delegated punishment options significantly increase cooperation compared to an environment without punishment, the evidence on their relative performance is mixed. Nosenzo and Sefton (2014) and Carpenter et al. (2012) find that cooperation and efficiency are significantly higher under decentralized sanctions. Other studies find no significant differences between the two regimes (O'Gorman et al., 2009; Fischer et al., 2016; Ambrus and Greiner, 2019) while Andreoni and Gee (2012) report significantly higher cooperation and efficiency under centralized sanctions. In our setting, the expected sanction for non-compliance is never lower, and typically higher, under delegated punishment than under peer punishment: In DP

---

[6]Some hypotheses may differ slightly from the one preregistered for exposition purposes.

regimes, punishable groups members lose their entire deposit of 6 tokens with certainty. In PP regimes, punishable group members can escape all or part of that loss if one or more of their peers in the group decide not to pay to become eligible for learning in the Monitoring Stage. This leads to our second hypothesis.

Hypothesis 2 *(Punishment)*: Contributions are a) highest in treatments with a delegated punishment regime and b) lowest in treatments with a self-punishment regime.

With the third and fourth hypothesis, we turn to the effect of monitoring on contributions. In previous experiments, contributions have been shown to increase with the level of information group members have about each other, even when information is not actionable (e.g. Sell and Wilson, 1991; Kreitmair, 2015). In our setting, all SP treatments that rely on self-punishment of non-compliance share the feature that information about other group members' individual contributions and compliance status is not actionable. By design this information is not available in the baseline, but available in those treatments that combine SP with DM and CM. We therefore predict in Hypothesis 3 that the monitoring opportunity has specific behavioral consequences.

Hypothesis 3 *(Monitoring)*: In treatments with self-punishment only, contributions are higher when participants can receive information about their group members' individual contributions than when they only receive aggregated information about the group average contribution.

At the same time, when information is costly, participants willingness to pay for monitoring is usually limited (Kurzban and DeScioli, 2008; Tasch and Houser, 2018; Liu et al., 2022), even when information is a prerequisite for punishment (Goeschl and Jarke, 2017; DeAngelo and Gee, 2020; Nicklisch et al., 2021). In our setting, CM implies monitoring effort in every round and leads to other group members' contributions and compliance status becoming available to all group members whenever monitoring is successful. Under

DM, by contrast, the provision of monitoring effort by participants is a matter of individual choice at the cost of 1 token in order to become eligible for learning this information.

As a result, non-compliance is more likely to be detected and to be detected by more group members in treatments with a CM regime compared to those with a DM regime, leading to a stronger deterrence effect of punishment. This is captured in Hypothesis 4.

Hypothesis 4 *(Decentralized vs. centralized monitoring)*: Contributions are higher under centralized monitoring than under decentralized monitoring.

3.2.2 Efficiency

Monitoring and punishment institutions are expected to have positive effects on contributions. These effects need not translate into positive effects on efficiency, as measured by participants' net payoffs, for two reasons: Actual punishments destroy earnings; and the monitoring-and-punishment institution is costly.

On the first point, the evidence in the experimental literature is mixed and depends on how the punishment options are design. Net gains in payoffs as a result of a punishment option (*e.g.,* Fehr and Gächter, 2000) and net losses (*e.g.,* Nikiforakis (2008); Denant-Boemont et al. (2007)) have both been observed, with the different findings driven in no small part by antisocial punishment (Herrmann et al., 2008). Our design rules out antisocial punishment: Participants are only allowed to punish non-compliant group members. As a result, we expect the destructive effect of actual punishment on earnings to be limited, giving rise to Hypothesis 5.

Hypothesis 5 *(Baseline vs. treatments)*: Net payoffs are never lower in the six treatment conditions than in the baseline.

In addition, in the absence of the destructive effects of antisocial punishment, we expect the more stringent punishment mechanisms to lead to higher net payoffs. As the

expected costs of non-compliance increases, free-riders will be deterred more often and - consequently - punishment is less likely to occur. This reasoning underpins the interpretation of previous experiments that have shown that delegated punishment can lead to higher average net payoffs than peer punishment (Andreoni and Gee, 2012) or no punishment (Tyran and Feld, 2006). This leads to Hypothesis 6.

Hypothesis 6 *(Punishment)*: Net payoffs are highest with delegated punishment.

The costliness of the monitoring-and-punishment institution arises in our design because becoming eligible for learning the results of successful monitoring requires costly investment. This cost is particularly relevant in treatments with CM. There, all participants are forced to pay for becoming eligible for learning. CM therefore involves (weakly) higher institutional costs than DM.[7] Hence, for the same level of cooperation, net payoffs can be higher under decentralized than under centralized monitoring. However, we posit that coordination among group members is unlikely to arise as communication is not possible in our setting. Moreover, monitoring costs are low relative to the gains from cooperation.[8] Hence, we expect the relative increase in cooperation under centralized monitoring to lead to higher net payoffs than under decentralized monitoring. This leads to our final hypothesis.

Hypothesis 7 *(Monitoring)*: Net payoffs are higher under centralized than under decentralized monitoring.

---

[7]This is especially the case with delegated punishment, as under DM, a single group member paying for monitoring is sufficient for punishment to occur. Consequently with DM, DP can achieve the same level of enforcement as CM at a quarter of the cost. We discuss the validity of this design feature in the discussion section.

[8]Note that to compensate the total monitoring cost of centralized monitoring (4 tokens), each group member needs to contribute an additional token to the Group Account ((0.4*4-1)*4.

4 Results

We first provide tests of the four main hypotheses about differences in contributions between treatments. We then move on to the remaining three hypothesis that make predictions about the relative efficiency of each of our institutions, as measured by participants' average net earnings, and participants' willingness to pay to monitor their group members under decentralized monitoring across enforcement mechanisms. As pre-registered, we excluded from our analyses groups for which at least one participant dropped out of the study from the first round. If a participant dropped out after round 1, we also excluded observations for all rounds after which the participant dropped out for that group. Thus, all the analyses reported below include 932 participants in 233 groups.

4.1 Main Results

4.1.1 Contributions

Figure 1 displays the evolution of the average group contributions across rounds, by treatment. For orientation, the figure also shows, as a dashed red line, the contribution level (8 tokens) that makes participants compliant.

Looking at the SP-NM baseline of no monitoring and self-punishment (in black), we repeat the typical finding for the public goods game without a punishment mechanism that average contributions are well above the Nash equilibrium prediction of zero for a selfish rational player and declining towards the final round.[9] Clearly, for the baseline, the presence of a compliance threshold at 8 tokens does not appear to be a factor in driving cooperation: Average contributions stay consistently at around 6 tokens. Second, for the other treatments, we also find that contributions clearly exceed zero in any given round. Visually, a first pattern appears in that SP treatments (orange) exhibit a decline towards the final round while PP (green) and DP (blue) treatments exhibit an increase. Third, only the DP treatments see average contributions meet or exceed the compliance threshold

---

[9]The phenomenon that the decline in contributions across rounds in the public goods game without punishment is less pronounced in online environments than in laboratory environments has been observed in some studies Diederich et al. (*e.g.,* 2016), but not others (*e.g.,* Arechar et al., 2018).

at some point during the repeated interaction. These patterns already foreshadow some of the results of a more rigorous analysis of the data.
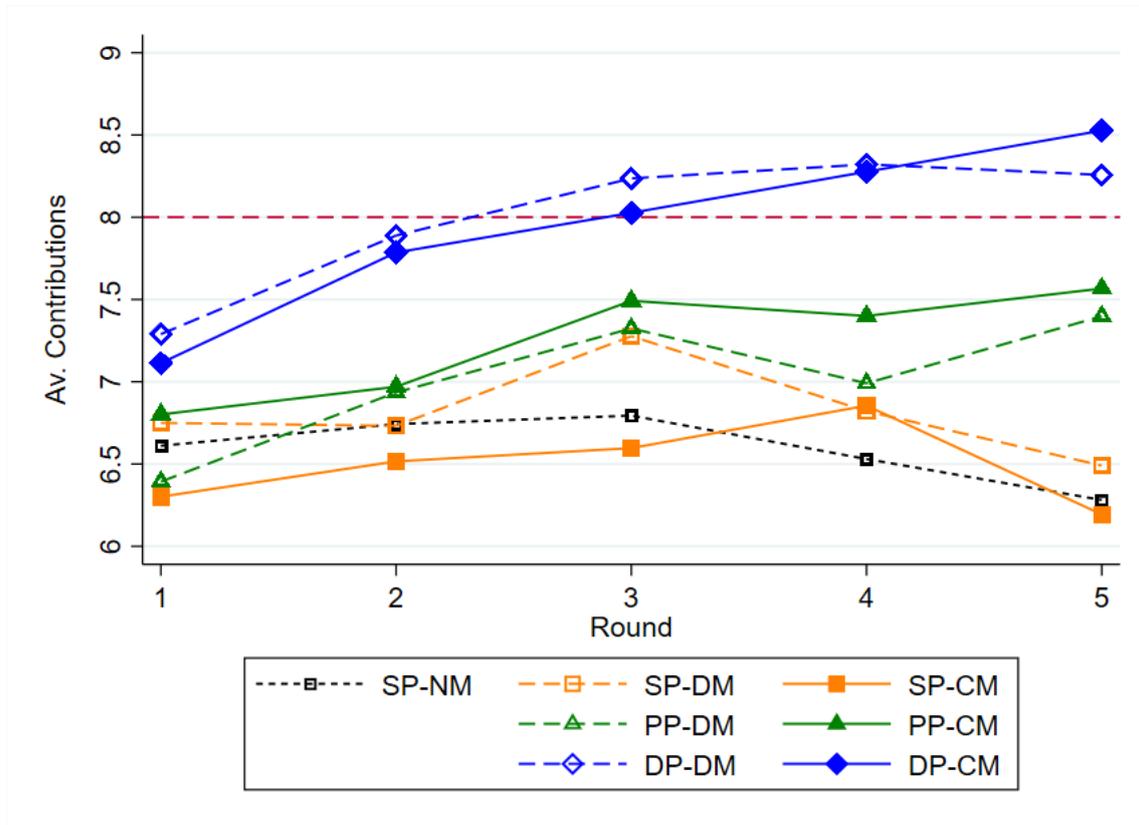

Figure 1: Average group contribution over time by treatments.

To compare how each of our institution performed relative to the baseline, Table 2 displays the mean contributions across all rounds separately for each treatment. We find that, compared to SP-NM, average contributions are significantly higher in treatments with a DP regime, both when combined with a CM (Fligner-Policello test:[10] $p < 0.001$) and DM regime (FP: $p < 0.001$).[11] In treatments with a PP regime, average contributions are also significantly higher than in the baseline when combined with a CM regime (FP: p=0.033), but not when combined with a DM regime (FP: $p = 0.353$). In treatments with an SP regime, combining this weak form of sanctioning with a CM (FP: $p = 0.774$) or a DM regime (FP: $p = 0.514$) fails to raise average contributions relative to the baseline. Contrary to Sell and Wilson (1991) or Kreitmair (2015), we therefore find that mutual monitoring alone is not capable of fostering cooperation in our experiment when a credible

---

[10]FP, hereafter.

[11]All reported non-parametric tests are two-sided.

enforcement mechanism is absent.

Table 2: Summary statistics.

	N Part.	N groups	Contributions		Net payoffs	
			Mean	(s.e.)	Mean	(s.e.)
SP-NM	144	36	6.57	(0.227)	19.48	(0.174)
SP-DM	120	30	6.79	(0.226)	20.21***	(0.200)
PP-DM	112	28	6.95	(0.270)	19.91	(0.249)
DP-DM	152	38	7.92***	(0.153)	20.44***	(0.183)
SP-CM	136	34	6.51	(0.241)	19.58	(0.165)
PP-CM	136	34	7.24*	(0.232)	19.71	(0.223)
DP-CM	132	33	7.81***	(0.197)	19.82*	(0.254)

Note: Average of the main outcome variables over the five rounds by treatment. Standard errors in parentheses. Stars indicates significant differences from the baseline (SP-NM). *p<0.05, **p<0.01, ***p<0.001.

To probe these results further and add robustness, we use a random-effects regression approach to estimate the treatment effects on average group contributions. Our dependent variable is the average contribution of group i in round t. The explanatory variables comprise a constant, a treatment dummy as well as round fixed effects. Column (1) in Table 3 displays the GLS coefficients for a comparison with the baseline (the reference category) and reaffirms the results of the non-parametric tests. Coefficients for treatments that feature PP and DP regimes are consistently positive, regardless of the monitoring regime. The coefficients for treatments that feature an SP regime are positive for DM and negative for CM regimes, but not significantly. This finding is in line with Hypothesis 1, leading to Result 1.

Result 1 *(Baseline vs. treatments)*: Contributions in the six treatment conditions with some monitoring-and-punishment institution are not lower than contributions in a baseline condition with no monitoring and self-punishment.

In hindsight, Hypothesis 1 turns out to have been on the conservative side regarding the predicted treatment effects. The results reported in Table 3 show that contributions in our experiment were significantly higher than under the baseline for all monitoring regimes when combined with delegated or peer punishment regimes, except when peer punishment is combined with decentralized monitoring (PP-DM). This underlines that contributions are reliably higher when a credible enforcement mechanism is present compared to the baseline.

Table 3: Effect of the enforcement mechanisms and monitoring institutions on group average contributions.

Dependent var:	Average contributions for group i in round t	
	(1)	(2)
SP-NM	*Ref.*	—
SP-DM	0.222	*Ref.*
	(0.312)	
SP-CM	-0.072	-0.293
	(0.302)	(0.312)
PP-DM	0.384	0.159
	(0.317)	(0.326)
PP-CM	0.658*	0.438
	(0.302)	(0.312)
DP-DM	1.367***	1.144***
	(0.294)	(0.316)
DP*CM	1.294***	1.076***
	(0.306)	(0.316)
Round FE	X	X
Constant	6.201***	6.347***
	(0.217)	(0.234)
Obs.	1085	916
Cluster	233	197

Note: Table 3 displays the GLS coefficients. Standard errors in parentheses. Stars indicates significant differences from the reference category. *p<0.05, **p<0.01, ***p<0.001.

Visual inspection of Figure 1 suggested the presence of performance differences between the three punishment regimes in combination with the two monitoring regimes. Column (2) of Table 3 provides GLS estimates of the joint effect of combining monitoring and punishment regimes on group average contributions relative to the SP-DM treatment,

thus excluding the no-monitoring baseline and controlling for round fixed-effects. We find a clear ranking of average contributions by punishment regime: Average contributions are significantly higher in treatments with a DP regime compared to those with an SP regime, both under centralized (DP-CM vs. SP-CM: Chi2(1): $p < 0.001$) and decentralized (DP-DM vs. SP-DM: $p < 0.001$) monitoring. Likewise, average contributions are significantly higher in treatments with a DP regime compared to those with a PP regime, both under centralized (DP-CM vs. PP-CM: Chi2(1): $p = 0.037$) and decentralized (DP-DM vs. PE-DM: $p = 0.001$) monitoring. Treatments with a DP regime therefore consistently outperform those with PP and SP alternatives.

Attempting to rank the PP and SP regimes, we find that average contributions under a PP regime are significantly higher than those under an SP regime when monitoring is centralized (PP-CM vs. SP-CM: Chi2(1): $p = 0.016$), but not when it is decentralized (PP-DM vs. SP-DM: $p = 0.626$). Peer punishment therefore does not generally outrank self-punishment in our experiment, leading to Result 2.

Result 2 *(Punishment)*: Comparing punishment regimes, delegated punishment induces the highest contributions irrespective of the monitoring regime, followed by peer punishment with centralized monitoring. Peer punishment with decentralized monitoring and regimes with self-punishment rank last in terms of average contributions.

Pooling treatments by punishment regime, groups contribute on average 1.27 tokens more under delegated punishment than under self-punishment and 0.8 tokens more than groups under peer punishment ($p < 0.001$ in both cases), as a random-effects GLS estimation shows.[12] Jointly, these results underline that among our institutional alternatives, delegated punishment is best suited to supporting cooperation, irrespective of the monitoring regime.

For a better understanding of the role of monitoring in curtailing free-riding, we return

---

[12]See column (1) of Table A-1 in Appendix A-2.

to column (1) of Table 3, specifically the coefficients associated with groups operating in the absence of a credible punishment regime. Relative to the baseline, we find no significant effect on average contributions of having a CM ($p = 0.812$) or DM regime ($p = 0.478$) present in an SP treatment. This is in line with the results of our non-parametric tests that monitoring alone cannot foster compliance in the absence of a credible enforcement mechanism, but runs counter to Hypothesis 3. This leads to Result 3.

Result 3 *(Monitoring)*: Monitoring alone does not improve contributions relative to the baseline in the absence of a credible enforcement mechanism.

To rank decentralized and centralized monitoring, we examine – for each punishment regime – whether average contributions differ significantly between DM and CM regimes. We find no significant differences for any of the enforcement mechanism (SP-DM vs. SP-CM: $p = 0.348$; PP-DM vs. PP-CM: Chi(1): $p = 0.379$; DP-DM vs. DP-CM: Chi(1): $p = 0.820$). This finding is supported by a random-effect regression comparing the average relative effect of DM and CM regimes across all three punishment regimes ($p = 0.532$).[13] This provides the answer to Hypothesis 4.

Result 4 *(Decentralized vs. centralized monitoring)*: We find no significant differences in the effect of decentralized and centralized monitoring on average contributions.

In summary, we found that delegated punishment not only performs best in enhancing contributions relative to peer- and self-punishment, but also raises average contributions above the compliance threshold, both under centralized and decentralized monitoring.[14] In addition, we found that the presence of a monitoring device alone cannot foster contributions beyond the baseline level in the absence of a credible enforcement mechanism. Altogether, our findings provide support for the idea that enforcement matters more than

---

[13]See column (2) of Table A-1 in Appendix A-2.

[14]Using two-sided Wilcoxon signed-rank tests, we cannot reject the hypothesis that average contributions across all rounds are equal to 8 in both DP-CM (p=0.851) and DP-DM (p=0.919).

monitoring in explaining variations in contributions.

4.1.2 Efficiency

Delegated punishment emerged as the superior institutional arrangement for raising contributions (Result 2). This comparative success needs to be compared with the resources that delegated punishment requires to perform such an improvement and compared with the resources required by different alternatives, possibly affecting their relative ranking. Groups' net payoffs are the essential metric for comparing efficiency across monitoring and punishment variations.

Figure 2 displays, for all treatments, the evolution of net payoffs over time.
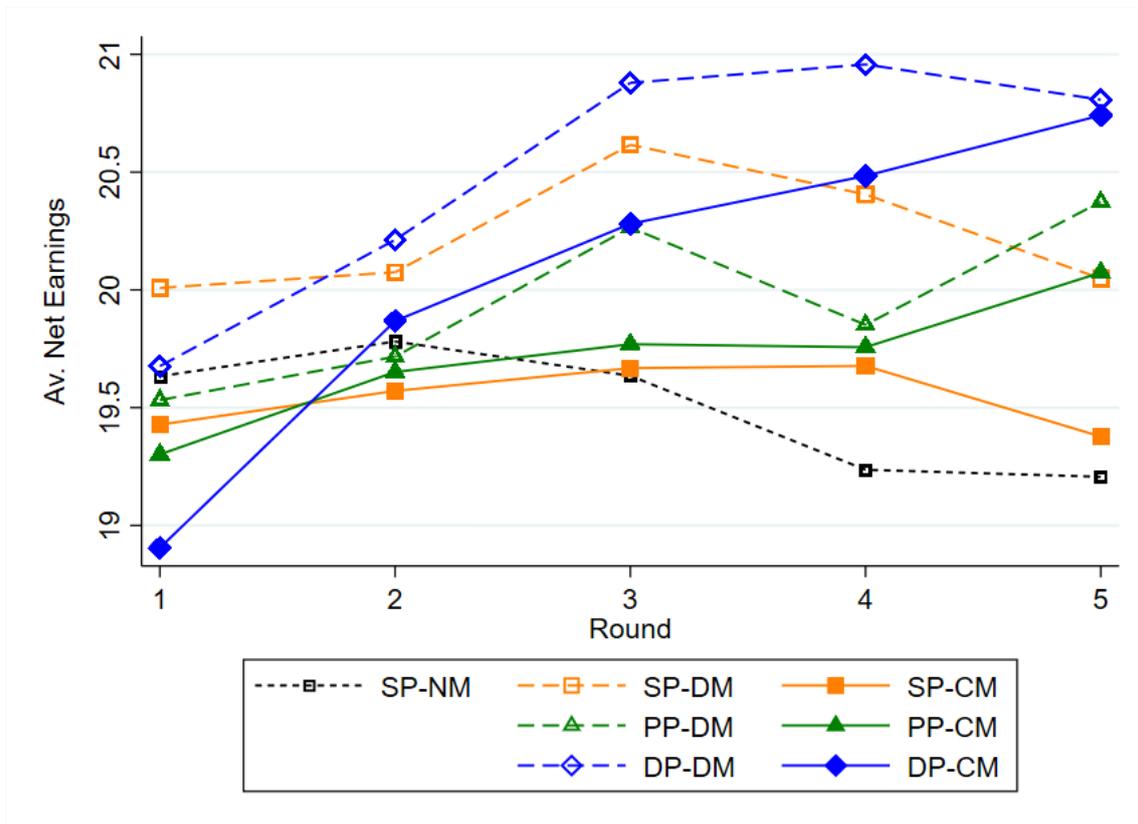

Figure 2: Distribution of net payoffs by treatment.

Visually, we see that the average net payoffs of most institutional alternatives start close together at around 19.5 tokens in round 1 before diverging somewhat across rounds. By round 5, the net payoff associated with the baseline is the lowest, at around 19.2 tokens, and the DP-DM and DP-CM treatments are the highest at around 20.8 tokens.

These observations are underlined by Table 2, which reports average net payoffs by treatment. Comparing these payoff, we find that relative to the SP-NM baseline, net payoffs are significantly higher in SP-DM (FP: $p = 0.002$), DP-DM (FP: $p < 0.001$), and DP-CM (FP: $p = 0.047$). There are no significant differences in net payoffs between the baseline and SP-CM (FP: $p = 0.760$), PP-DM (FP: $p = 0.133$) and PP-CM (FP: $p = 0.265$). Given that centralized monitoring is costly (by design) and that willingness to pay for monitoring is positive under decentralized monitoring (see section 4.2), our results suggest that the increase in contributions generated by the presence of a credible enforcement mechanism is sufficient to compensate for the cost of monitoring. Altogether, these results are in line with Hypothesis 5 and show that none of institutional alternatives performs worse than the baseline.

Result 5 *(Baseline vs. treatments)*: Net payoffs in the six treatment conditions with some monitoring-and-punishment institution are not lower than net payoffs in the baseline with no monitoring and self-punishment.

As in the previous section, we also report on the role the punishment and monitoring regimes each have in determining efficiency. First, pooling across monitoring institutions, we find that delegated, peer- and self-punishment regimes do not induce significant differences in efficiency terms. This is the results of a random-effect GLS regression.[15] This suggests that, contrary to Hypothesis 6, the increase in contributions generated by a credible punishment mechanism (see Result 2) does not guarantee sufficient gains to compensate for the additional cost of running the institution.

Result 6 *(Punishment)*: We find no significant difference in net payoffs between any of the enforcement mechanism (does not support H6).

Conversely, pooling across enforcement mechanisms, we find – using the same approach

---

[15]See column (3) of Table A-1 in Appendix A-2.

– that net payoffs are significantly higher under decentralized than under centralized monitoring ($p = 0.004$).[16] This finding runs counter to Hypothesis 7, which speculated that the (weakly) higher cost of centralized monitoring would be more than offset by gains in contribution. Instead, it turns out that the lower cost of decentralized monitoring more than offset the losses from smaller contributions, compared to the centralized monitoring regime.

Result 7 *(Monitoring)*: Net payoffs are significantly higher under decentralized monitoring than under centralized monitoring (does not support H7).

We investigate the mechanism behind this result in the next section.

4.2 Willingness to Pay for Monitoring

By design of the experiment, decentralized monitoring is at most as costly as centralized monitoring and never more so. This feature is favorable to the efficiency of treatments with a DM regime and renders Result 7 intuitive. At the same time, the design of decentralized monitoring is also unfavorable towards efficiency because it provides only very limited opportunities for group members to coordinate on monitoring. Result 7 makes clear that groups in treatments with a DM regime were not adversely affected by the unfavorable factors, however.

Figure 3 displays the average group willingness-to-pay over time under decentralized monitoring, by punishment regime.

---

[16]See the coefficients of a a random-effect regression reported in column (4) of Table A-1 in Appendix A-2.
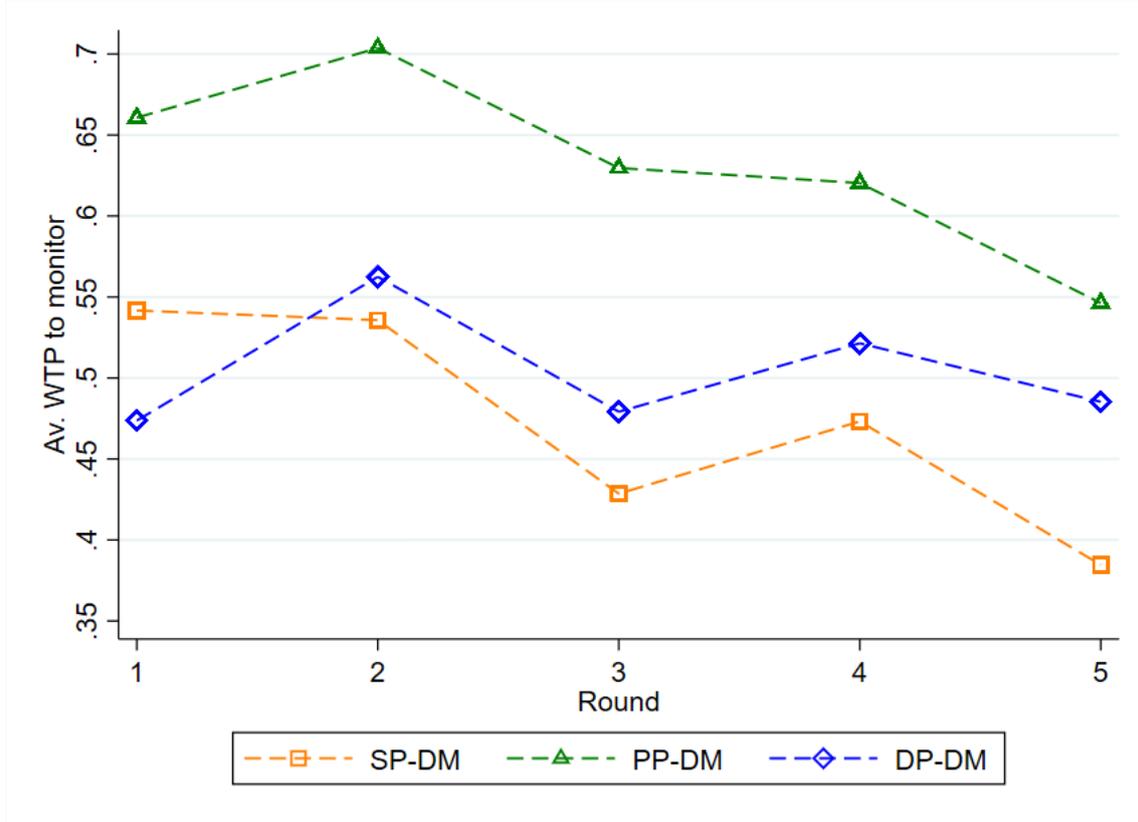
Figure 3: Average group willingness-to-pay to monitor over time under decentralized monitoring, by enforcement mechanisms.

First, note that in all treatment, participants spend significantly less than 1 token on monitoring under decentralized monitoring (Wilcoxon signed-ranked tests: $p < 0.001$ in all three treatments): On average, group member expend around 0.55 tokens and at most 0.7 tokens per round, instead of 1 token per round under centralized monitoring. Decentralized monitoring therefore incurs significantly lower cost of running the punishment institution.

Second, note that participants' willingness to pay to become eligible for receiving information about their group members' behavior under a DM regime varied systematically with the punishment mechanism. The visual evidence suggests, and non-parametric tests show, that participants' willingness-to-pay is higher under a peer-punishment regime than under a self-punishment regime (0.46 vs. 0.64; FP: $p < 0.001$). The willingness-to-pay is also higher under a peer-punishment regime than under a delegated punishment regime (0.64 vs. 0.50; FP: $p < 0.001$). There is no significant difference, however, in willingness-

to-pay between self-punishment and delegated punishment (0.46 vs. 0.50; FP: $p = 0.344$), even though the instrumental value of monitoring in the latter is much higher than in the former. The complementarity between willingness-to-pay to become eligible for learning and the ability to punish at the individual level harnesses the innate demand to punish free-riding group members. The relatively low willingness-to-pay observed under delegated punishment, by contrast, is consistent with the public goods nature of paying to become eligible for learning: Only one person in the group needs to pay those costs to enable the punishment of successfully monitored group members found to have been non-compliant, but at least one person has to pay. As a result, a substantial share of participants will be tempted to free-ride on their group members for monitoring purposes. We summarize these observations in a final result.

Result 8 *(WTP Monitoring)*: Under peer-monitoring, participants are willing to pay to monitor their group members, and significantly more so with peer-punishment than with any other punishment mechanism, but less than under centralized monitoring.

This result helps explain how treatments with a decentralized monitoring regime achieve higher net payoffs than those with a centralized monitoring regime while achieving the same level of contributions.

5 Discussion and Conclusion

Curtailing free-riding is the main function of a number of institutional features that surround social dilemmas. Institutions that support punishment, either by peers or through a delegated mechanism, have been studied extensively and the question of how monitoring of group members' behavior contributes to curtailing free-riding has attracted scholars' attention more recently. In this paper, we built on this rich existing literature to study how jointly varying the monitoring and punishment regimes of an institutional solution to a social dilemma affect outcomes. This allows us to capture more of the richness of

real-world institutional variation than was previously possible – and to expand the set of those situations to which the experimental literature can speak.

Our pre-registered experimental design examined, as a baseline, how group members perform in a public goods game with no monitoring of other group members' individual contribution behavior and access to self-punishment only. Against this baseline, it assessed the performance of six alternative institutional set-ups, systematically manipulating which of three punishment mechanisms (self-, peer- or delegated punishment) and two (imperfect) monitoring regimes (centralized vs. decentralized) would be applied to a group of participants in an online setting.

Overall, we found evidence for the primacy of the punishment regime for curtailing free-riding: Variations in the punishment mechanisms are responsible for most of the observed variation in contributions. Variations in the monitoring institutions are contributing relatively little. More specifically, the evidence is in line with our hypotheses that contributions are higher with peer-punishment than with self-punishment, and that delegated punishment outperforms both alternatives. In contrast with the literature on information provision in public good games, we did not find evidence that – in the absence of a credible punishment mechanism – providing individual information about group members' contributions raises contributions compared to providing only aggregated information. In contrast with an earlier experiment that varied how monitoring is organized, we found no significant differences in contributions between centralized and decentralized monitoring under peer-punishment. This invariance of contributions to the organization of monitoring extended to delegated punishment. Under efficiency considerations, however, it mattered: Decentralized monitoring achieved higher net payoffs than centralized monitoring as participants spent less money on monitoring.

Taken together, our findings suggest to the institutional designer that her first and foremost consideration should be how punishment is organized, with the organization of monitoring a secondary concern. The consideration of punishment regimes should include

a recognition of the significant ability of delegated punishment to limit free-riding. At the same time, the institutional design should also heed the observation that more centralized institutions do not have to be preferable *per se.* In our experiment, for example, decentralized monitoring did not lead to less compliance than centralized monitoring, providing degrees of freedom to the institutional designer.

Our experimental design incorporates a number of critical choices that can capture some real-world settings better than alternative ones, providing opportunities for future research. First, we designed decentralized monitoring to be (weakly) less expensive than centralized monitoring. This captures, for instance, the context of international environmental agreements. There the creation of centralized institutions typically requires the creation of new administrative bodies such as secretariats or independent agencies. Agreement on these institutions is often hard to reach and if so, only after long delays and at a considerable cost of time and money. By contrast, decentralized monitoring often already exists, such as remote sensing for scientific reasons, and therefore benefits from pre-existing infrastructure. In other settings, and in the long run, centralized institutions may well be cheaper than mutual monitoring by each party involved. Whether our results would be replicated in such circumstances is an open question.

Second, it is worth noting that contributions in our baseline are overall higher than what is typically observed in public good game experiments. This could be driven by the presence of the compliance threshold. Previous papers have found that such thresholds can function as a focal point in terms of expected cooperation and increase contributions compared to a typical public good game, even in the absence of any credible punishment mechanism. Higher contributions in the baseline condition limit the potential treatment effect that can be observed. As a result, our results may well be capturing a lower bound of the mechanisms under study.

Third, in contrast to some of the existing literature, our joint varying of monitoring and punishment regimes imposed an institutional structure exogenously. This constituted a natural first step towards answering the question of joint variation of regimes. An

obvious next step is to examine which of the six institutions would have been chosen endogenously by our participants, and how such an exogenous choice may affect free-riding and efficiency.

References

Ambrus, A. and Greiner, B. (2012). Imperfect public monitoring with costly punishment: An experimental study. *American Economic Review*, 102(7):3317–32.

Ambrus, A. and Greiner, B. (2019). Individual, dictator, and democratic punishment in public good games with perfect and imperfect observability. *Journal of Public Economics*, 178:104053.

Andreoni, J. and Gee, L. K. (2012). Gun for hire: Delegated enforcement and peer punishment in public goods provision. *Journal of Public Economics*, 96(11-12):1036–1046.

Arechar, A. A., Gächter, S., and Molleman, L. (2018). Conducting interactive experiments online. *Experimental economics*, 21:99–131.

Baldassarri, D. and Grossman, G. (2011). Centralized sanctioning and legitimate authority promote cooperation in humans. *Proceedings of the National Academy of Sciences*, 108(27):11023–11027.

Barron, K. and Nurminen, T. (2020). Nudging cooperation in public goods provision. *Journal of Behavioral and Experimental Economics*, 88:101542.

Carpenter, J., Kariv, S., and Schotter, A. (2012). Network architecture, cooperation and punishment in public good experiments. *Review of Economic Design*, 16:93–118.

Casari, M. (2007). Emergence of endogenous legal institutions: Property rights and community governance in the italian alps. *The Journal of Economic History*, 67(1):191–226.

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

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental economics*, 14(1):47–83.

Cherry, T. L. and McEvoy, D. M. (2013). Enforcing compliance with environmental agreements in the absence of strong institutions: An experimental analysis. *Environmental and Resource Economics*, 54(1):63–77.

Cherry, T. L. and McEvoy, D. M. (2017). Refundable deposits as enforcement mechanisms in cooperative agreements: Experimental evidence with uncertainty and non-deterrent sanctions.

Dannenberg, A. and Gallier, C. (2020). The choice of institutions to solve cooperation problems: a survey of experimental research. *Experimental Economics*, 23(3):716–749.

DeAngelo, G. and Gee, L. K. (2020). Peers or police?: The effect of choice and type of

monitoring in the provision of public goods. *Games and Economic Behavior*, 123:210–227.

Denant-Boemont, L., Masclet, D., and Noussair, C. N. (2007). Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. *Economic theory*, 33(1):145–167.

Diederich, J., Goeschl, T., and Waichman, I. (2016). Group size and the (in) efficiency of pure public good provision. *European Economic Review*, 85:272–287.

Egas, M. and Riedl, A. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 275(1637):871–878.

Falkinger, J., Fehr, E., Gächter, S., and Winter-Ebmer, R. (2000). A simple mechanism for the efficient provision of public goods: Experimental evidence. *American economic review*, 91(1):247–264.

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

Fehr, E. and Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868):137–140.

Fischer, S., Grechenig, K., and Meier, N. (2016). Monopolizing sanctioning power under noise eliminates perverse punishment but does not increase cooperation. *Frontiers in behavioral neuroscience*, 10:180.

Fudenberg, D. and Pathak, P. A. (2010). Unobserved punishment supports cooperation. *Journal of Public Economics*, 94(1-2):78–86.

Gerber, A. and Wichardt, P. C. (2009). Providing public goods in the absence of strong institutions. *Journal of Public Economics*, 93(3-4):429–439.

Goeschl, T. and Jarke, J. (2017). Trust, but verify? monitoring, inspection costs, and opportunism under limited observability. *Journal of Economic Behavior & Organization*, 142:320–330.

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.

Gross, J., Méder, Z. Z., Okamoto-Barth, S., and Riedl, A. (2016). Building the leviathan–voluntary centralisation of punishment power sustains cooperation in humans. *Scientific reports*, 6(1):20767.

Guillen, P., Schwieren, C., and Staffiero, G. (2007). Why feed the leviathan? *Public Choice*, 130:115–128.

Herrmann, B., Thoni, C., and Gachter, S. (2008). Antisocial punishment across societies. *Science*, 319(5868):1362–1367.

Kreitmair, U. W. (2015). Voluntary disclosure of contributions: an experimental study on nonmandatory approaches for improving public good provision. *Ecology and Society*, 20(4).

Kurzban, R. and DeScioli, P. (2008). Reciprocity in groups: information-seeking in a public goods game. *European Journal of Social Psychology*, 38(1):139–158.

Liu, J., Sonntag, A., and Zizzo, D. J. (2022). Information defaults in repeated public

good provision. *Journal of Economic Behavior & Organization*, 197:356–369.

Nicklisch, A., Grechenig, K., and Thöni, C. (2016). Information-sensitive leviathans. *Journal of Public Economics*, 144:1–13.

Nicklisch, A., Putterman, L., and Thöni, C. (2021). Trigger-happy or precisionist? on demand for monitoring in peer-based public goods provision. *Journal of Public Economics*, 200:104429.

Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics*, 92(1-2):91–112.

Nikiforakis, N. and Engelmann, D. (2011). Altruistic punishment and the threat of feuds. *Journal of Economic Behavior & Organization*, 78(3):319–332.

Nosenzo, D. and Sefton, M. (2014). Promoting cooperation: the distribution of reward and punishment power.

O'Gorman, R., Henrich, J., and Van Vugt, M. (2009). Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 276(1655):323–329.

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.

Page, T., Putterman, L., and Unel, B. (2005). Voluntary association in public goods experiments: Reciprocity, mimicry and efficiency. *The Economic Journal*, 115(506):1032–1053.

Piza, E. L., Welsh, B. C., Farrington, D. P., and Thomas, A. L. (2019). Cctv surveillance for crime prevention: A 40-year systematic review with meta-analysis. *Criminology & public policy*, 18(1):135–159.

Rustagi, D., Engel, S., and Kosfeld, M. (2010). Conditional cooperation and costly monitoring explain success in forest commons management. *science*, 330(6006):961–965.

Sell, J. and Wilson, R. K. (1991). Levels of information and contributions to public goods. *Social Forces*, 70(1):107–124.

Sutter, M., Haigner, S., and Kocher, M. G. (2010). Choosing the carrot or the stick? endogenous institutional choice in social dilemma situations. *The Review of Economic Studies*, 77(4):1540–1566.

Tasch, W. and Houser, D. (2018). Social preferences and social curiosity.

Traulsen, A., Röhl, T., and Milinski, M. (2012). An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proceedings of the Royal Society B: Biological Sciences*, 279(1743):3716–3721.

Tyran, J.-R. and Feld, L. P. (2006). Achieving compliance when legal sanctions are non-deterrent. *scandinavian Journal of Economics*, 108(1):135–156.

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

Appendix A   Additional Analyses and Tables

A-1   Overview of the data

Figure A-1 displays the distribution of contributions across treatments averaged across all rounds. Figure A-2 displays the distribution of net payoffs across treatments averaged across all rounds.
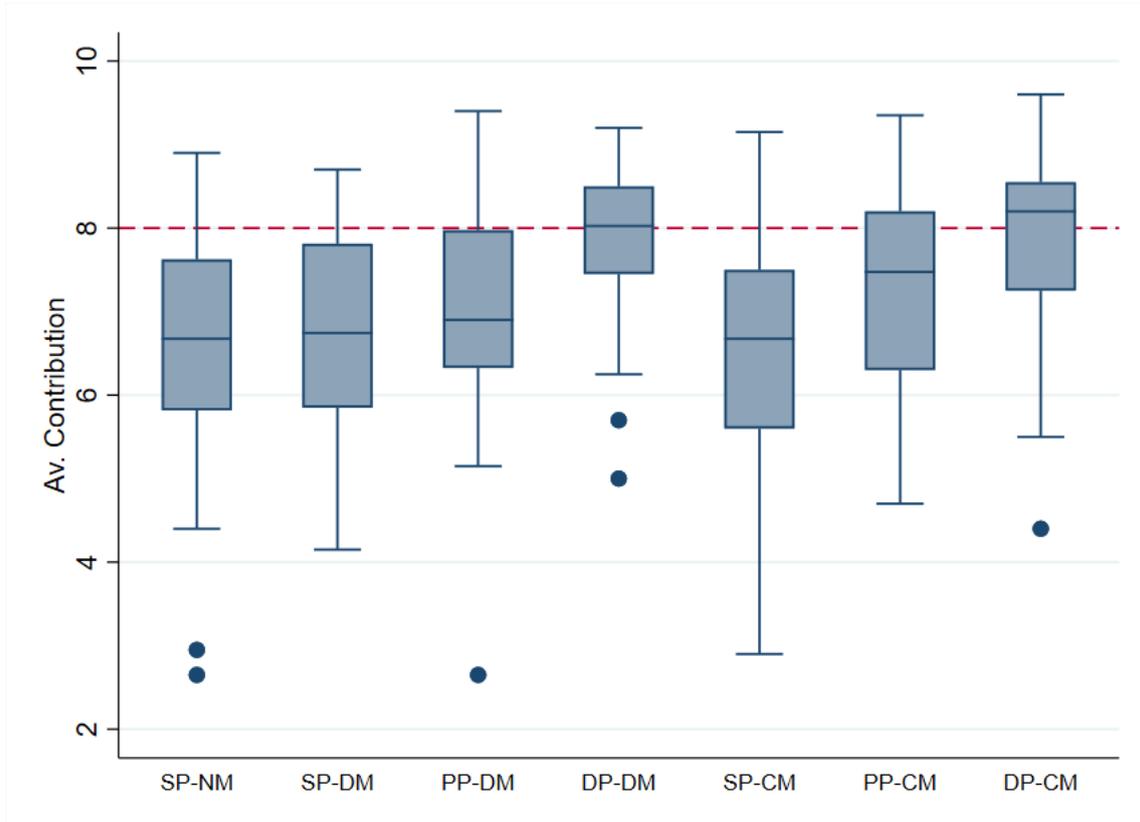

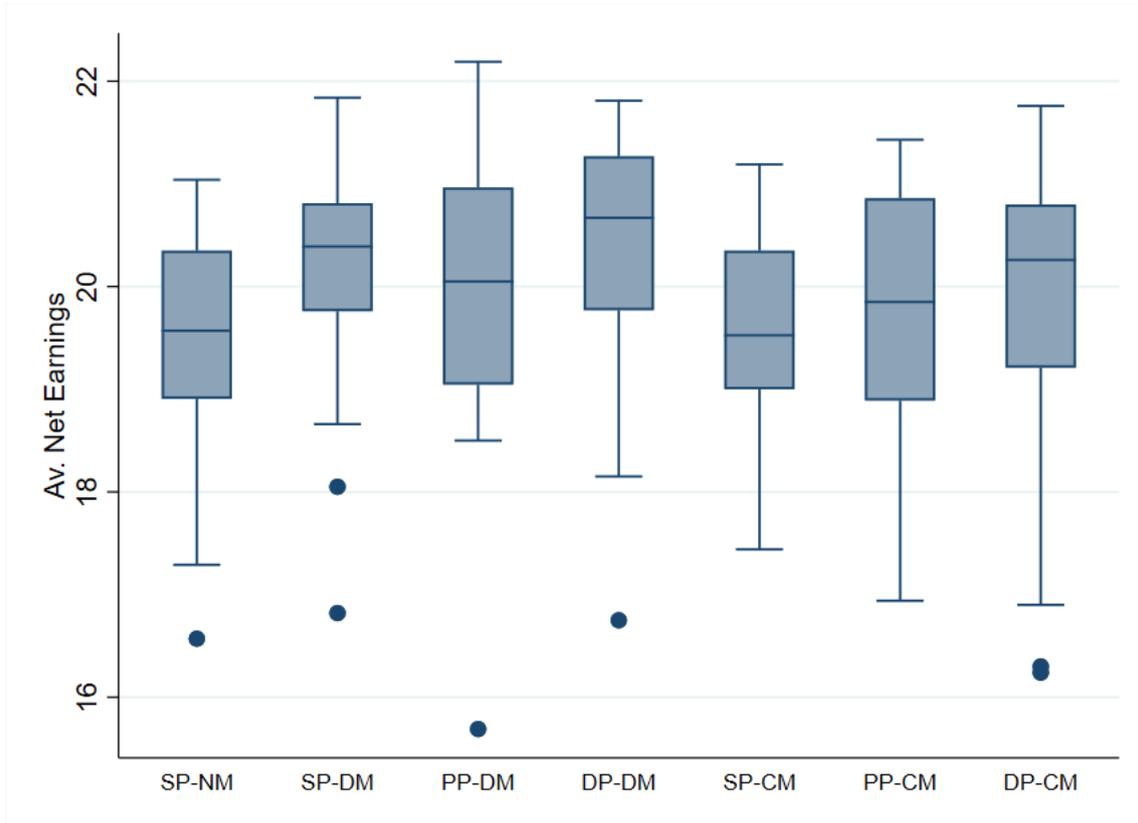Figure A-1: Distribution of contributions by treatment.

Figure A-2: Distribution of net payoffs by treatment.

A-2   Main Results

To examine the relative importance of enforcement mechanisms vs. monitoring regime on contributions further, we estimated the aggregated effect of our enforcement mechanisms and monitoring regime on group average contributions separately. We first conducted a random-effect regression using a dummy variable for each of our enforcement mechanisms as our main independent variables (excluding self-enforcement as the reference category), controlling for round fixed-effects. The results are displayed in column (1). In addition, we conducted a random-effect regression using a dummy variable equal to 1 if the group experienced centralized monitoring and 0 otherwise as our main independent variable, controlling for round fixed-effects. The results are displayed in column (2).

Table A-1: Effect of the enforcement mechanisms and monitoring institutions on group average contributions.

Dependent var:	Average contributions for group i in round t		Average net payoffs for group i in round t	
	(1)	(2)	(3)	(4)
SE	*Ref.*	–	*Ref.*	–
PE	0.467*	–	-0.074	–
	(0.220)		(0.0216)	
3E	1.268***	–	0.350	–
	(0.214)		(0.210)	
PM	–	*Ref.*	–	*Ref.*
3M	–	-0.101	–	-0.493***
		(0.192)		(0.171)
Round FE	X	X	X	X
Constant	6.192***	6.847***	19.37***	19.72***
	(0.164)	(0.148)	(0.166)	(0.140)
Obs.	916	916	916	916
Cluster	197	197	197	197

Note: Table A-1 displays the GLS coefficients. Standard errors in parentheses. Stars indicates significant differences from the *Ref.* category. *p<0.05, **p<0.01, ***p<0.001.

A-2.1 Willingness to Punish

In this section, we investigate the proportion of participants who have been punished after having been found to be non-compliant in each treatment. Figure A-3 displays participants' average punishment rate over time in each treatment.[17]

---

[17]Because punishment is automated in treatments with third-party enforcement, we excluded thoese treatments from the Figure.
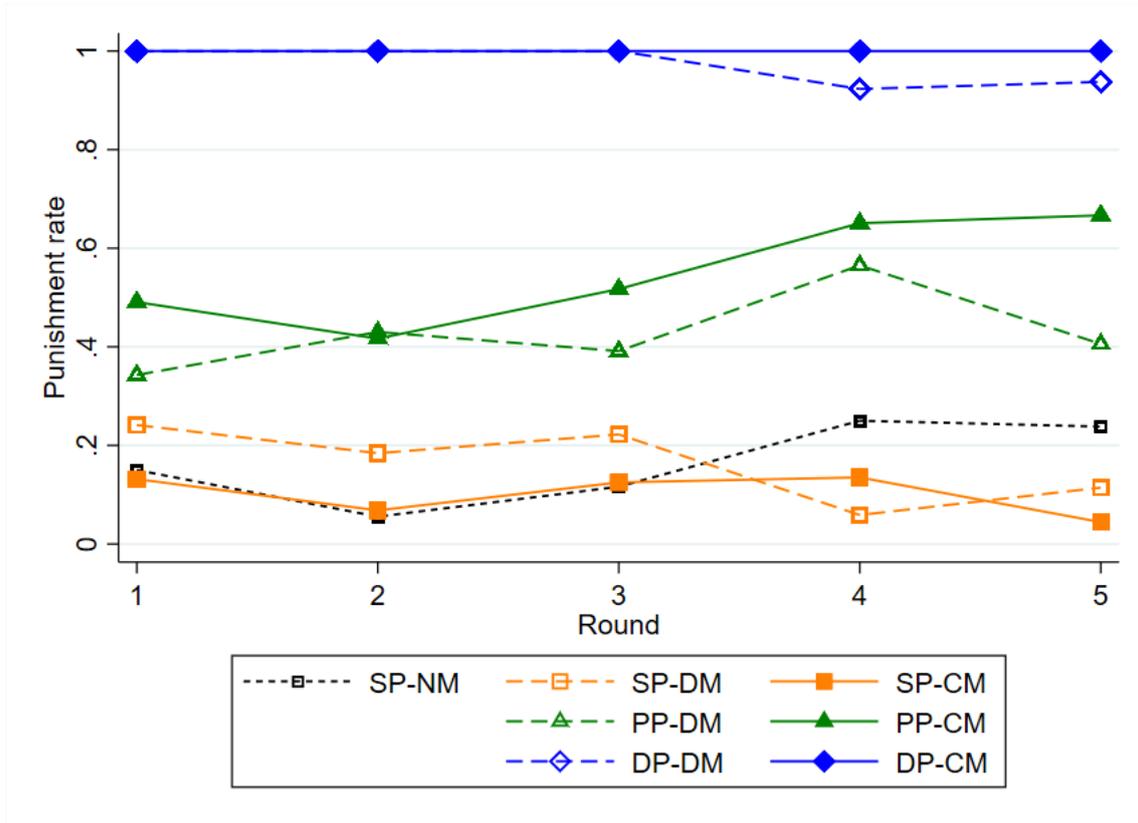
Figure A-3: Average group willingness-to-punish non-compliant group members over time, by treatment.

Figure A-3 shows that the punishment rate is higher with peer-enforcement than with self-enforcement, regardless of the monitoring institution. However, the Figure also shows that participants who experienced peer-enforcement do not sanction to the full extent as the punishment rate is still way below 1. This is especially surprising under centralized monitoring as monitoring is already paid for and sanctioning is therefore costless at this point. Figure A-3 also shows that a positive fraction of participants are willing to forfeit their own deposit in the self-enforcement treatment, regardless of the monitoring institution.

We find that the punishment rate is significantly higher with peer enforcement than with self-enforcement (0.47 (0.019) vs. 0.14 (0.015); FP: $p < 0.001$) but significantly lower than 1 ($p = XX$). In addition, we find no significant difference in punishment rate between the different monitoring regime with self-enforcement ($p < 0.001$), nor with peer-enforcement ($p = XX$). In summary, the punishment rate is higher with peer-enforcement than with self-enforcement regardless of the monitoring institution, but participants do not punish to the full extent.

A-3   Summary statistics

A-4   Balance Check

Table A-2 provides summary statistics of the demographic variables for participants in the seven treatment conditions. Using one-way ANOVAs, we find no significant dif-

ference between treatments in the percentage of female (F(1,930)=1.36, p=0.231), age (F(46,858)=1.04, p=0.395), perceived clarity of the instructions (F(3,901)=0.36, p=0.784), nor in the percentage of participants who indicated French as their native language (F(1,903)=0.34, p=0.562).

Table A-2: Summary statistics of demographic variables.

	N Part.	N Group	Female	Age	English	Clarity
SE-NM	144	36	61.11%	27.55 (0.236)	38.13%	2.27 (0.021)
SE-PM	120	30	54.17%	26.99 (0.247)	39.66%	2.27 (0.025)
PE-PM	112	28	50.89%	27.11 (0.277)	35.14%	2.22 (0.024)
3E-PM	152	38	50.66%	26.79 (0.232)	43.24%	2.21 (0.022)
SE-3M	136	34	50%	26.91 (0.231)	36.09%	2.12 (0.024)
PE-3M	136	34	52.21%	27.53 (0.293)	42.42%	2.25 (0.021)
3E-3M	132	33	54.55%	26.06 (0.216)	40.48%	2.22 (0.022)
			$p = 0.231$	$p = 0.395$	$p = 0.562$	$p = 0.784$

Note: Female is an indicator that is equal to 1 if the participant is female. English is an indicator that is equal to 1 if the participant answered "Yes" to the question: "Is English the language you use most at home?". Risk is the participant's answer to the question: "Are you a person who is generally willing to take risks, or do you try to avoid taking risks on a scale from 0 (completely unwilling to take risks) to 10 (completely willing to take risks)? Clarity is the participant's answer to the question: "How would you rate the clarity of the questions?" (Answer possibilities: 0 Extremely unclear, 1 Somewhat unclear, 2 Somewhat clear, 3 Extremely clear). Standard deviations are in parentheses. P-values are for one-way ANOVAs.

Appendix B   Instructions

[COMMON TO ALL TREATMENTS]

General instructions (1)

Participants will be interacting with each other.

It is therefore necessary and in the interest of all participants to detect whether individuals have dropped out from the study.

To do this, we have set a reasonable, but fixed amount of time to complete each page.

On some pages, you will make **decisions**.

The computer program will consider that you dropped out of the study if you fail to submit a decision in the allocated time.

In this case, the remaining pages of the study will submit automatically and you will not receive any payment.

If you cannot commit to complete the study in the allocated time, please exit the study now.

If you are ready to start the study, press the 'Next' button.

Next

General instructions (2)

During this study, you will interact with other Prolific users.

Hence, you may sometimes have to wait for others. When this happens, this will be indicated to you by a waiting screen similar to the one below.

Please wait

Waiting for the other participants.
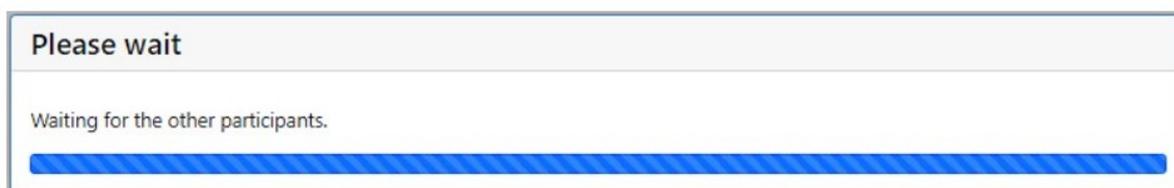
Next

Welcome!

Thank you for taking part in this study.

You have been randomly assigned to a group of 4 participants.

During this study you will have to make some decisions. Your earnings will depend on your decisions, as well as the decisions of your fellow group members.

During the study we shall not speak of pounds, but of tokens. Your entire earnings will be calculated in tokens. At the end of the study, the total amount of tokens you have earned will be converted into pounds according to the following exchange rate:

10 tokens = £1

This study is composed of four stages: a Contribution Stage, an Information Stage, an Enforcement Stage, and a Feedback Stage.

Next

Procedure

You and the 3 other participants in your group will play the same game for 5 rounds.

Your group members will remain the same across all 5 rounds.

Each group member will be given one of the following ID: P1, P2, P3, or P4.

These IDs will remain the same for all participants across all 5 rounds.

At the end of the study, the computer program will randomly select one of the 5 rounds for payment. Your total earnings for this round will be paid to you in addition to your participation fee.

Since all rounds are equally likely to be chosen, you should make your decision in each round as if it will be the round-that-counts.

Next

39

Accounts

In this study, each group member has three accounts: two private ones and one shared with the group.

At the beginning, that is at the Contribution Stage, each group member starts with the following account holdings:

- A private Active Account with 10 tokens that can be used during the Contribution Stage.
- A private Passive Account with a deposit of 6 tokens.
- A shared Group Account with 0 tokens.

ACTIVE ACCOUNT	PASSIVE ACCOUNT	GROUP ACCOUNT
10 Tokens	6 Tokens	0 Tokens

In the later Stages, account holdings may increase or decrease depending on your decisions and those of the other members in your group.

Next

Stage 1: Contribution Stage

At the beginning of the Contribution Stage, you (and each of your three group members) will receive 10 tokens placed in your private Active Account (**your endowment**) and 6 tokens placed in your private Passive Account (**your deposit**).

YOUR TASK

Your task is to decide how many of the 10 tokens from your private Active Account you want to contribute to the shared Group Account.

At the end of the Contribution Stage, every 1 token contributed to the Group Account is multiplied by 1.6 and the product is divided equally among the four group members. So, for every 1 token contributed to the Group Account during the Contribution Stage, everyone in the group receives **0.40 token** in their Active Account at the end of the Contribution Stage.

You may choose to contribute any number of tokens to the Group Account you like. We will call your contribution **high** if you contribute **8 tokens or more**, and **low** if you contribute **7 tokens or less.**

Any tokens that you do not contribute to the Group Account stay in your Active Account.

Each other member of your group will have to make the same decision.

40

CONTRIBUTION STAGE EARNINGS

At the end of the Contribution Stage, your Active Account contains earnings that are calculated as follows:

$$\text{Contribution Stage Earnings} = \text{Your endowment of 10 in your Active Account}$$
$$- \text{ Your contribution to the Group Account}$$
$$+ 1.6 \times \frac{(\text{Tokens in the Group Account})}{4}$$

Contribution Stage earnings are calculated by this formula for all four group members.

<div align="right">

Next

</div>

[NO MONITORING (SP-NM) - Monitoring Stage instructions]

Stage 2: Information Stage

Time remaining on this page: **1:58**

In the Information Stage, you will be reminded of your own contribution decision in the Contribution Stage. You will also see your group total contribution to the Group Account.

The screen looks like this:

GROUP ACCOUNT:
[Your group total contribution to the Group Account]

Players	Contributions
P1 (You)	[Your contribution to the Group Account]
P2	???
P3	???
P4	???

<div align="right">

Next

</div>

Stage 2: Information Stage

Time remaining on this page: **1:46**

At the beginning of the Information Stage, you (and each of your three group members) receive one additional token in your private Active Account.

You can choose to use this token to pay an access fee for the opportunity to receive additional information about the individual contribution decisions of group members. The additional information you receive is determined by the following process: A computer program will roll a three-sided die for each group member. If the die lands on 1 or 2, the group member's contribution will be made available <u>to those who paid the access fee</u>. If the die lands on 3, the group member's contribution will not be made available. You will be informed whether your individual contribution decision was made available. The screen looks like this:

GROUP ACCOUNT:
[Your group total contribution to the Group Account]

You chose to pay the access fee.

Players	Die roll	Contributions
P1 (You)	⚁	[Your contribution to the Group Account]
P2	⚀	[P2's contribution to the Group Account]
P3	⚂	???
P4	⚁	[P4's contribution to the Group Account]

If you choose to keep this token and pay <u>no</u> access fee, then you will be reminded of your own contribution decision in the Contribution Stage. You will also see the total contribution to the Group Account. The screen looks like this:

GROUP ACCOUNT:
[Your group total contribution to the Group Account]

You chose NOT to pay the access fee.

Players	Die roll	Contributions
P1 (You)	⚅(2)	[Your contribution to the Group Account]
P2	⚀(1)	???
P3	⚂(3)	???
P4	⚁(2)	???

Next

[CENTRALIZED MONITORING (SP-CM, PP-CM, DP-CM) - Monitoring Stage instructions]

Stage 2: Information Stage

Time remaining on this page: **1:53**

At the beginning of the Information Stage, you (and each of your three group members) receive one additional token in your private Active Account.

You and your fellow group members automatically use this token to pay an access fee for the opportunity to receive additional information about the individual contribution decisions of group members. The additional information you receive is determined by the following process: A computer program will roll a three-sided die for each group member. If the die lands on 1 or 2, the group member's contribution will be made available to <u>all group members</u>. If the die lands on 3, the group member's contribution will not be made available. You will be informed whether your individual contribution decision was made available.

In addition, you will be reminded of your own contribution decision in the Contribution Stage. You will also see the total contribution to the Group Account. The screen looks like this:

GROUP ACCOUNT:
[Your group total contribution to the Group Account]

Players	Die roll	Contributions
P1 (You)	[die: 3]	[Your contribution to the Group Account]
P2	[die: 1]	[P2's contribution to the Group Account]
P3	[die: 4]	???
P4	[die: 2]	[P4's contribution to the Group Account]

Next

[SELF PUNISHMENT (SP-NM, SP-DM, SP-CM) - Punishment Stage instructions]

Stage 3: Enforcement Stage

Time remaining on this page: **1:59**

In the Enforcement Stage, you may be able to forfeit your own deposit. To be able to do so, the following condition needs to be fulfilled:

1. Your contribution was low.

If you decide to forfeit your deposit, all 6 tokens will be deducted from your private Passive Account.

Each other member of your group will have to make the same decision.

Next

[PEER PUNISHMENT (PP-NM, PP-DM, PP-CM) - Punishment Stage instructions]

Stage 3: Enforcement Stage

Time remaining on this page: **1:55**

In the Enforcement Stage, you may be able to reduce other group members' deposits. To be able to do so, the following conditions need to be fulfilled:

1. You have paid the access fee at the Information Stage.
2. The group member's contribution was made available to you.
3. The group member's contribution was low.

If you decide to reduce another group member's deposit, 2 tokens will be deducted from their private Passive Account.

To you, there is no cost associated with this decision.

Each other member of your group will make the same decision.

Next

[DELEGATED PUNISHMENT (DP-NM, DP-DM, DP-CM) - Punishment Stage instructions]

Stage 3: Enforcement Stage

Time remaining on this page: **1:54**

In the Enforcement Stage, a computer program will forfeit a group member's deposit if the following conditions are fulfilled:

1. At least one group member has paid the access fee at the Information Stage.
2. The group member's contribution was made available.
3. The group member's contribution was low.

If the computer program forfeits a group member's deposit, all 6 tokens will be deducted from their private Passive Account.
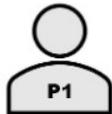
Next

45

[(SP-NM) - Feedback Stage instructions + Summary]

Stage 4: Feedback Stage

In the Feedback Stage, you receive information about how many tokens (out of 6) are left on each group member's Passive Account. The screen looks like this:

| You have *[number of tokens]* tokens left in your Passive Account. | P2 has *[number of tokens]* tokens left in their Passive Account. | P3 has *[number of tokens]* tokens left in their Passive Account. | P4 has *[number of tokens]* tokens left in their Passive Account. |

EARNINGS

In addition, you will be informed about your total earnings from the four stages.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$

+

PASSIVE ACCOUNT

6 tokens if your contribution was high;
OR 6 tokens if your contribution was low and you do not forfeit your deposit;
OR 0 token if your contribution was low and you forfeit your deposit.

Next

46

Summary of the instructions

- You have been matched with 3 other participants and you will play the game with the same participants for 5 rounds.
- In the Contribution Stage, you decide how many tokens from your Active Account to contribute to the Group Account.
  - Your contribution will be considered high if you contribute 8 tokens or more to the Group Account, otherwise your contribution will be considered low.
  - Your contribution stage earnings are computed as follows:
    Your endowment of 10 tokens - Your contribution to the Group Account + 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
- In the Information Stage, you will be informed about your group total contribution to the Group Account.
- In the Enforcement Stage, you will decide whether to forfeit your 6-token deposit if your contribution was low.

Your total earnings are calculated as follows:

<table>
<tr><td>

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$

</td><td>

+

</td><td>

PASSIVE ACCOUNT

6 tokens if your contribution was high;
<u>OR</u> 6 tokens if your contribution was low and you <u>do not</u> forfeit your deposit;
<u>OR</u> 0 token if your contribution was low and you forfeit your deposit.

</td></tr>
</table>

Next

Stage 4: Feedback Stage

Time remaining on this page: **1:51**

In the Feedback Stage, you receive information about how many tokens (out of 6) are left on each group member's Passive Account. The screen looks like this:

P1	P2	P3	P4
You have *[number of tokens]* tokens left in your Passive Account.	P2 has *[number of tokens]* tokens left in their Passive Account.	P3 has *[number of tokens]* tokens left in their Passive Account.	P4 has *[number of tokens]* tokens left in their Passive Account.

EARNINGS

In addition, you will be informed about your total earnings from the four stages.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token <u>if you pay</u> the access fee

+

PASSIVE ACCOUNT

6 tokens if your contribution was high;
<u>OR</u> 6 tokens if your contribution was low and you <u>do not</u> forfeit your deposit;
<u>OR</u> 0 token if your contribution was low and you forfeit your deposit.

48

Summary of the instructions

- You have been matched with 3 other participants and you will play the game with the same participants for 5 rounds.
- In the Contribution Stage, you decide how many tokens from your Active Account to contribute to the Group Account.
  - Your contribution will be considered high if you contribute 8 tokens or more to the Group Account, otherwise your contribution will be considered low.
  - Your contribution stage earnings are computed as follows:
    Your endowment of 10 tokens - Your contribution to the Group Account + 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
- In the Information Stage, you will be informed about your group total contribution to the Group Account. In addition, you can pay a 1 token access fee to receive information about the contribution decision of some of your group members.
  - If you decide to pay the access fee, whether a group member's contribution will be made available depends on the outcome of a die roll.
- In the Enforcement Stage, you will decide whether to forfeit the 6-token deposit if your contribution was low.

Your total earnings are calculated as follows:

<table>
<tr><td>

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token <u>if you pay</u> the access fee

</td><td>

+

</td><td>

PASSIVE ACCOUNT

6 tokens if your contribution was high;
<u>OR</u> 6 tokens if your contribution was low and you <u>do not</u> forfeit your deposit;
<u>OR</u> 0 token if your contribution was low and you forfeit your deposit.

</td></tr>
</table>

[(SP-CM) - Feedback Stage instructions + Summary]

Stage 4: Feedback Stage

In the Feedback Stage, you receive information about how many tokens (out of 6) are left on each group member's Passive Account. The screen looks like this:

P1	P2	P3	P4
You have *[number of tokens]* tokens left in your Passive Account.	P2 has *[number of tokens]* tokens left in their Passive Account.	P3 has *[number of tokens]* tokens left in their Passive Account.	P4 has *[number of tokens]* tokens left in their Passive Account.

EARNINGS

In addition, you will be informed about your total earnings from the four stages.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
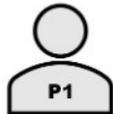- 1 token access fee

+

PASSIVE ACCOUNT

6 tokens if your contribution was high;
OR 6 tokens if your contribution was low and you do not forfeit your deposit;
OR 0 token if your contribution was low and you forfeit your deposit.

50

Summary of the instructions

- You have been matched with 3 other participants and you will play the game with the same participants for 5 rounds.
- In the Contribution Stage, you decide how many tokens from your Active Account to contribute to the Group Account.
  - Your contribution will be considered high if you contribute 8 tokens or more to the Group Account, otherwise your contribution will be considered low.
  - Your contribution stage earnings are computed as follows:
    Your endowment of 10 tokens - Your contribution to the Group Account + 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
- In the Information Stage, you will be informed about your group total contribution to the Group Account. In addition, you will pay a 1 token access fee to receive information about the contribution decision of some of your group members.
  - Whether a group member's contribution will be made available depends on the outcome of a die roll.
- In the Enforcement Stage, you will decide whether to forfeit the 6-token deposit if your contribution was low.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT		PASSIVE ACCOUNT
+ Your Contribution Stage Endowment of 10 tokens - Your contribution to the Group Account + 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$ + Your Information Stage Endowment of 1 token - 1 token access fee	**+**	6 tokens if your contribution was high; OR 6 tokens if your contribution was low and you do not forfeit your deposit; OR 0 token if your contribution was low and you forfeit your deposit.

51

[(PP-DM) - Feedback Stage instructions + Summary]

Stage 4: Feedback Stage

In the Feedback Stage, you receive information about how many tokens (out of 6) are left on each group member's Passive Account. The screen looks like this:

| You have *[number of tokens]* tokens left in your Passive Account. | P2 has *[number of tokens]* tokens left in their Passive Account. | P3 has *[number of tokens]* tokens left in their Passive Account. | P4 has *[number of tokens]* tokens left in their Passive Account. |

EARNINGS

In addition, you will be informed about your total earnings from the four stages.

Your total earnings are calculated as follows:

EARNINGS

In addition, you will be informed about your total earnings from the four stages.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token if you pay the access fee

+

PASSIVE ACCOUNT

6 tokens if your contribution was high;
OR 6 tokens if your contribution was not found low;
OR 6 tokens if your contribution was found low and **no one** reduced your deposit;
OR 4 tokens if your contribution was found low and **1 group member** reduced your deposit;
OR 2 tokens if your contribution was found low and **2 group members** reduced your deposit;
OR 0 token if your contribution was found low and **3 group members** reduced your deposit.

Next

52

Summary of the instructions

- You have been matched with 3 other participants and you will play the game with the same participants for 5 rounds.
- In the Contribution Stage, you decide how many tokens from your Active Account to contribute to the Group Account.
  - Your contribution will be considered high if you contribute 8 tokens or more to the Group Account, otherwise your contribution will be considered low.
  - Your contribution stage earnings are computed as follows:
    Your endowment of 10 tokens - Your contribution to the Group Account + 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
- In the Information Stage, you will be informed about your group total contribution to the Group Account. In addition, you can pay a 1 token access fee to receive information about the contribution decision of some of your group members.
  - If you decide to pay the access fee, whether a group member's contribution will be made available depends on the outcome of a die roll.
- In the Enforcement Stage, you will decide whether to reduce the 6-token deposit of group members whose contribution was found to be low.

Your total earnings are calculated as follows:

<table>
<tr><td>

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token <u>if you pay</u> the access fee

</td><td>

+

</td><td>

PASSIVE ACCOUNT

6 tokens if your contribution was high;
<u>OR</u> 6 tokens if your contribution was <u>not</u> found low;
<u>OR</u> 6 tokens if your contribution was found low and **no one** reduced your deposit;
<u>OR</u> 4 tokens if your contribution was found low and **1 group member** reduced your deposit;
<u>OR</u> 2 tokens if your contribution was found low and **2 group members** reduced your deposit;
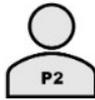<u>OR</u> 0 token if your contribution was found low and **3 group members** reduced your deposit.

</td></tr>
</table>
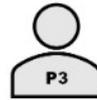
Next

Stage 4: Feedback Stage

In the Feedback Stage, you receive information about how many tokens (out of 6) are left on each group member's Passive Account. The screen looks like this:
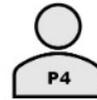
P1	P2	P3	P4
You have *[number of tokens]* tokens left in your Passive Account.	P2 has *[number of tokens]* tokens left in their Passive Account.	P3 has *[number of tokens]* tokens left in their Passive Account.	P4 has *[number of tokens]* tokens left in their Passive Account.

EARNINGS

In addition, you will be informed about your total earnings from the four stages.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token access fee

+

PASSIVE ACCOUNT

6 tokens, if your contribution was high;
OR 6 tokens if your contribution was <u>not</u> found low;
OR 6 tokens if your contribution was found low and **no one** reduced your deposit;
OR 4 tokens if your contribution was found low and **1 group member** reduced your deposit;
OR 2 tokens if your contribution was found low and **2 group members** reduced your deposit;
OR 0 token if your contribution was found low and **3 group members** reduced your deposit.

Next

Summary of the instructions

- You have been matched with 3 other participants and you will play the game with the same participants for 5 rounds.
- In the Contribution Stage, you decide how many tokens from your Active Account to contribute to the Group Account.
  - Your contribution will be considered high if you contribute 8 tokens or more to the Group Account, otherwise your contribution will be considered low.
  - Your contribution stage earnings are computed as follows:
    Your endowment of 10 tokens - Your contribution to the Group Account + 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
- In the Information Stage, you will be informed about your group total contribution to the Group Account. In addition, you will pay a 1 token access fee to receive information about the contribution decision of some of your group members.
  - Whether a group member's contribution will be made available depends on the outcome of a die roll.
- In the Enforcement Stage, you will decide whether to reduce the 6-token deposit of group members whose contribution was found to be low.

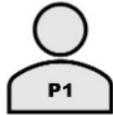Your total earnings are calculated as follows:

<table>
<tr>
<td>

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token access fee

</td>
<td>

+

</td>
<td>

PASSIVE ACCOUNT

6 tokens, if your contribution was high;
OR 6 tokens if your contribution was not found low;
OR 6 tokens if your contribution was found low and **no one** reduced your deposit;
OR 4 tokens if your contribution was found low and **1 group member** reduced your deposit;
OR 2 tokens if your contribution was found low and **2 group members** reduced your deposit;
OR 0 token if your contribution was found low and **3 group members** reduced your deposit.

</td>
</tr>
</table>

Next

55

Stage 4: Feedback Stage

In the Feedback Stage, you receive information about how many tokens (out of 6) are left on each group member's Passive Account. The screen looks like this:

P1	P2	P3	P4
You have *[number of tokens]* tokens left in your Passive Account.	P2 has *[number of tokens]* tokens left in their Passive Account.	P3 has *[number of tokens]* tokens left in their Passive Account.	P4 has *[number of tokens]* tokens left in their Passive Account.

EARNINGS

In addition, you will be informed about your total earnings from the four stages.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ $1.6 \times \frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token <u>if you pay</u> the access fee

+

PASSIVE ACCOUNT

6 tokens if your contribution was high;
<u>OR</u> 6 tokens if your contribution was <u>not</u> found low;
<u>OR</u> 0 token if your contribution was found to be low by either 1, 2 or 3 group members.

56

Summary of the instructions

- You have been matched with 3 other participants and you will play the game with the same participants for 5 rounds.
- In the Contribution Stage, you decide how many tokens from your Active Account to contribute to the Group Account.
    - Your contribution will be considered high if you contribute 8 tokens or more to the Group Account, otherwise your contribution will be considered low.
    - Your contribution stage earnings are computed as follows:
    Your endowment of 10 tokens - Your contribution to the Group Account + 1.6 x $\frac{(\text{Tokens in the Group Account})}{4}$
- In the Information Stage, you will be informed about your group total contribution to the Group Account. In addition, you can pay a 1 token access fee to receive information about the contribution decision of some of your group members.
    - If you decide to pay the access fee, whether a group member's contribution will be made available depends on the outcome of a die roll.
- In the Enforcement Stage, the computer program will automatically forfeit the 6-token deposit of group members whose contribution was found to be low by at least one group member.
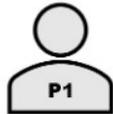
Your total earnings are calculated as follows:

<table>
<tr>
<td>

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{(\text{Tokens in the Group Account})}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token <u>if you pay</u> the access fee

</td>
<td>

+

</td>
<td>

PASSIVE ACCOUNT

6 tokens if your contribution was high;
<u>OR</u> 6 tokens if your contribution was <u>not</u> found low;
<u>OR</u> 0 token if your contribution was found to be low by either 1, 2 or 3 group members.

</td>
</tr>
</table>

57

Stage 4: Feedback Stage

Time remaining on this page: **1:54**

In the Feedback Stage, you receive information about how many tokens (out of 6) are left on each group member's Passive Account. The screen looks like this:

You have *[number of tokens]* tokens left in your Passive Account.

P2 has *[number of tokens]* tokens left in their Passive Account.

P3 has *[number of tokens]* tokens left in their Passive Account.

P4 has *[number of tokens]* tokens left in their Passive Account.

EARNINGS

In addition, you will be informed about your total earnings from the four stages.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT

+ 10 token Contribution Stage Endowment
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ 1 token Information Stage Endowment
- 1 token access fee

+

PASSIVE ACCOUNT

6 tokens if your contribution was high;
OR 6 tokens if your contribution was not found low;
OR 0 token if your contribution was found low.

Next

58

Summary of the instructions

- You have been matched with 3 other participants and you will play the game with the same participants for 5 rounds.
- In the Contribution Stage, you decide how many tokens from your Active Account to contribute to the Group Account.
  - Your contribution will be considered high if you contribute 8 tokens or more to the Group Account, otherwise your contribution will be considered low.
  - Your contribution stage earnings are computed as follows:
    Your endowment of 10 tokens - Your contribution to the Group Account + 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
- In the Information Stage, you will be informed about your group total contribution to the Group Account. In addition, you will pay a 1 token access fee to receive information about the contribution decision of some of your group members.
  - Whether a group member's contribution will be made available depends on the outcome of a die roll.
- In the Enforcement Stage, the computer program will automatically forfeit the 6-token deposit of group members whose contribution was found to be low by at least one group member.

Your total earnings are calculated as follows:

ACTIVE ACCOUNT

+ Your Contribution Stage Endowment of 10 tokens
- Your contribution to the Group Account
+ 1.6 x $\frac{\text{(Tokens in the Group Account)}}{4}$
+ Your Information Stage Endowment of 1 token
- 1 token access fee

+

PASSIVE ACCOUNT

6 tokens if your contribution was high;
OR 6 tokens if your contribution was <u>not</u> found low;
OR 0 token if your contribution was found low.

Comprehension Question

Time remaining on this page: **4:55**

You need to answer all the questions correctly before moving on to the next page.

If you need help, use the "REMINDER INSTRUCTIONS" button at the bottom of this page.

___

1) How many tokens do you earn in the Contribution Stage if you contribute 8 tokens to the Group Account and the three other members in your group also contribute 8 tokens each?

[   ] tokens.

___

2) How many tokens do you earn in the Contribution Stage if you contribute 0 token to the Group Account and the three other members in your group contribute 8 tokens each?

[   ] tokens.

5) My group members will remain the same throughout the 5 rounds.

○ True

○ False

[REMINDER INSTRUCTIONS] [CHECK ANSWERS]

[COMPREHENSION QUESTIONNAIRE (SP-NM)]

3) Check all correct statements.

☐ In the Information Stage, I will see my group total contribution to the Group Account.

☐ In the Information Stage, I will see the contribution of each other group members to the Group Account.

4) Check all correct statements.

☐ In the Enforcement Stage, I can forfeit my own deposit regardless of my contribution.

☐ In the Enforcement Stage, I can forfeit my own deposit only if my contribution was low.

☐ My group members will <u>not</u> see whether I decided to forfeit my own deposit, and I will <u>not</u> see whether my group members decided to forfeit their deposit.

☐ My group members will see whether I decided to forfeit my own deposit, and I will see whether my group members decided to forfeit their deposit.

[COMPREHENSION QUESTIONNAIRE (SP-DM)]

3) Check all correct statements.

☐ In the Information Stage, I will see my group total contribution to the Group Account.

☐ In the Information Stage, I will see the contribution of each other group members to the Group Account.

4) Check all correct statements.

- ☐ In the Enforcement Stage, I can forfeit my own deposit regardless of my contribution.

- ☐ In the Enforcement Stage, I can forfeit my own deposit only if my contribution was low.

- ☐ My group members will <u>not</u> see whether I decided to forfeit my own deposit, and I will <u>not</u> see whether my group members decided to forfeit their deposit.

- ☐ My group members will see whether I decided to forfeit my own deposit, and I will see whether my group members decided to forfeit their deposit.

[COMPREHENSION QUESTIONNAIRE (PP-DM)]

3) Check all correct statements about the Information Stage: Whether I see another group members' contribution the Group Account depends on...

- ☐ ... whether I paid the access fee or not.

- ☐ ... whether someone in the group paid the access fee or not.

- ☐ ... whether my contribution was high or low.

- ☐ ... the outcome of the die roll.

4) Check all correct statements.

- ☐ In the Enforcement Stage, I can reduce the deposit of group members whose contributions was found to be low, if I paid the access fee in the Information Stage.

- ☐ In the Enforcement Stage, I can reduce the deposit of group members whose contributions was found to be low, if at least one group member paid the access fee in the Information Stage.

- ☐ My group members will <u>not</u> see whether my deposit was reduced, and I will <u>not</u> see whether my group members' deposit was reduced.

- ☐ My group members will see whether my deposit was reduced, and I will see whether my group members' deposit was reduced.

[COMPREHENSION QUESTIONNAIRE (DP-DM)]

3) Check all correct statements about the Information Stage: Whether I see another group members' contribution the Group Account depends on...

- ☐ ... whether I paid the access fee or not.

- ☐ ... whether someone in the group paid the access fee or not.

- ☐ ... whether my contribution was high or low.

- ☐ ... the outcome of the die roll.

4) Check all correct statements.

☐ In the Enforcement Stage, the computer program will forfeit the deposit of group members whose contribution was found to be low, if I paid the access fee in the Information Stage.

☐ In the Enforcement Stage, the computer program will forfeit the deposit of group members whose contribution was found to be low, if at least one group member paid the access fee in the Information Stage.

☐ My group members will <u>not</u> see whether my deposit was forfeited, and I will <u>not</u> see whether my group members' deposit was forfeited.

☐ My group members will see whether my deposit was forfeited, and I will see whether my group members' deposit was forfeited.

[COMPREHENSION QUESTIONNAIRE (SP-CM)]

3) Check all correct statements about the Information Stage: Whether I see another group members' contribution the Group Account depends on...

☐ ... whether my contribution was high or low.

☐ ... whether the group member's contribution was high or low.

☐ ... the outcome of the die roll.

4) Check all correct statements.

☐ In the Enforcement Stage, I can forfeit my own deposit regardless of my contribution.

☐ In the Enforcement Stage, I can forfeit my own deposit only if my contribution was low.

☐ My group members will <u>not</u> see whether I decided to forfeit my own deposit, and I will <u>not</u> see whether my group members decided to forfeit their deposit.

☐ My group members will see whether I decided to forfeit my own deposit, and I will see whether my group members decided to forfeit their deposit.

[COMPREHENSION QUESTIONNAIRE (PP-CM)]

3) Check all correct statements about the Information Stage: Whether I see another group members' contribution the Group Account depends on...

☐ ... whether my contribution was high or low.

☐ ... whether the group member's contribution was high or low.

☐ ... the outcome of the die roll.

4) Check all correct statements.

☐ In the Enforcement Stage, I can reduce the deposit of group members whose contributions was found to be low.

☐ My group members will <u>not</u> see whether my deposit was reduced, and I will <u>not</u> see whether my group members' deposit was reduced.

☐ My group members will see whether my deposit was reduced, and I will see whether my group members' deposit was reduced.

[COMPREHENSION QUESTIONNAIRE (DP-CM)]

3) Check all correct statements about the Information Stage: Whether I see another group members' contribution the Group Account depends on...

☐ ... whether my contribution was high or low.

☐ ... whether the group member's contribution was high or low.

☐ ... the outcome of the die roll.

4) Check all correct statements.

☐ In the Enforcement Stage, the computer program will forfeit the deposit of group members whose contribution was low.

☐ My group members will <u>not</u> see whether my deposit was forfeited, and I will <u>not</u> see whether my group members' deposit was forfeited.

☐ My group members will see whether my deposit was forfeited, and I will see whether my group members' deposit was forfeited.

Stage 1: Round 1 out of 5

Time remaining on this page: **1:27**

Contribution Stage

In this stage, you have to decide how many tokens (from 0 to 10) to contribute to the group account.

I contribute [ ] **tokens** to the group account.

[REMIND ME OF THE GAME] [NEXT]

Stage 2: Round 1 out of 5

Time remaining on this page: **1:57**

Information Stage

GROUP ACCOUNT:
0

Players	Contributions
P1	???
P2	???
P3	???
P4 (You)	0

[Next]

[DECENTRALIZED MONITORING (DM) - Monitoring Stage]

64

Time remaining on this page: **0:57**

Information Stage

You received an additional token.

You can use this token to pay the access fee.

Do you want to pay the 1-token access fee?

○ No ○ Yes

If you do not pay the access fee, the additional token will be added to your earnings at the end of the round.

Next

Time remaining on this page: **1:43**

Information Stage

GROUP ACCOUNT:
20

You paid the access fee.

Players	Die roll	Contributions
P1	⚁	0
P2	⚂	???
P3 (You)	⚀	10 (Available to others)
P4	⚁	10

Next

Stage 2: Round 1 out of 5

Time remaining on this page: **1:31**

Information Stage

GROUP ACCOUNT:
20

You did not pay the access fee.

Players	Die roll	Contributions
(You)		0 *(Available to others)*
		???
		???
		???

Next

[CENTRALIZED MONITORING (DM) - Monitoring Stage]

Stage 2: Round 1 out of 5

Time remaining on this page: **0:56**

Information Stage

You received an additional token.

This token will be used to pay the access fee.

Next

66

Stage 2: Round 2 out of 5

Information Stage

GROUP ACCOUNT:
20

Players	Die roll	Contributions
P1		???
P2 (You)		0 *(Not available to others)*
P3		10
P4		10

Next

[SELF PUNISHMENT (SP) - Punishment Stage]

Stage 3: Round 1 out of 5

Time remaining on this page: **0:51**

Enforcement Stage

In this stage, you have to decide whether you want to forfeit your 6-token deposit.

Do you want to forfeit your own deposit?

○ No   ○ Yes

If choose not to forfeit your deposit, the 6 tokens will be added to your earnings at the end of the round.

Next

Reminder:

GROUP ACCOUNT:
0

Players	Contributions
P1	???
P2	???
P3	???
P4 (You)	0

68

Stage 3: Round 1 out of 5

Time remaining on this page: **1:51**

Enforcement Stage

In this stage, you have to decide for each group member whose contribution was found to be **low**, whether you want to reduce their 6-token deposit by 2 tokens or leave their deposit unchanged.

Players	Your Choice
P1	○ Leave deposit unchanged ○ Reduce deposit
P2	○ Leave deposit unchanged ○ Reduce deposit

Next

Reminder:

GROUP ACCOUNT:
20

You paid the access fee.

Players	Die roll	Contributions
P1	•	0
P2	•	0
P3 (You)	(die showing 5)	10 *(Not available to others)*
P4	•	10

69

[DELEGATED PUNISHMENT (DP) - Punishment Stage]

Stage 2: Round 2 out of 5

Time remaining on this page: **1:56**

Information Stage

GROUP ACCOUNT:
20

Players	Die roll	Contributions
P1		???
P2 (You)		0 *(Not available to others)*
P3		10
P4		10

Next

Stage 4: Round 1 out of 5

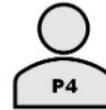Time remaining on this page: **0:58**

Feedback Stage

P1 has 6 tokens left in their Passive Account.

P2 has 0 tokens left in their Passive Account.

You have 0 tokens left in your Passive Account.

P4 has 6 tokens left in their Passive Account.

Your earnings for this round are 10.0 tokens.

Next

71