

Egberts, Alexander; Engel, Christoph; Fairfield, Joshua A. T.

**Working Paper**

## Out of sight is out of mind? Experimentally testing a gradually materializing public bad

Discussion Papers of the Max Planck Institute for Research on Collective Goods, No. 2024/16

**Provided in Cooperation with:**

Max Planck Institute for Research on Collective Goods

*Suggested Citation:* Egberts, Alexander; Engel, Christoph; Fairfield, Joshua A. T. (2024) : Out of sight is out of mind? Experimentally testing a gradually materializing public bad, Discussion Papers of the Max Planck Institute for Research on Collective Goods, No. 2024/16, Max Planck Institute for Research on Collective Goods, Bonn, <https://hdl.handle.net/21.11116/0000-0010-3CA2-4>

This Version is available at:

<https://hdl.handle.net/10419/307697>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



ALEXANDER EGBERTS  
CHRISTOPH ENGEL  
JOSHUA FAIRFIELD

Discussion Paper  
2024/16

OUT OF SIGHT IS  
OUT OF MIND?  
EXPERIMENTALLY  
TESTING A GRADUALLY  
MATERIALIZING PUBLIC  
BAD

**Out of Sight is Out of Mind?  
Experimentally Testing a Gradually Materializing Public Bad**

*Abstract*

*Many social ills can be modelled as a public bad. In such scenarios, private benefit is often immediate while the public damage takes some time to materialize. In this experiment, we investigate the behavioral effects caused by such delays in the realization of collective harm. By manipulating the weight with which the damages caused by group contributions are carried over to the next round, we alter the number of periods required for the social damage to fully unfold. We keep constant the economic consequences of contributions between treatments (by introducing a multiplier for the damage) and between periods (by deducting all unrealized harm at the end of the game) to avoid multiple equilibria. In a second treatment dimension, we isolate the cognitive challenges of this experiment by replacing human group-members with “computerized players” which perfectly copy each subject’s previous behavior. We find that participants’ behavior is less cooperative over time when harm is deferred into the future. Our results also suggest that the driving mechanism behind this effect is not insufficient anticipation, but the lack of having experienced the negative consequences of the public damage.*

Keywords: public bad; dynamically developing social harm; cognitive and motivational challenge; experiment

JEL: C91, D62, D91, H41, K24, K32

## 1. Introduction

**Research question.** A social dilemma is called a dilemma for a reason. If everybody affected by the incentive structure is exclusively interested in maximizing individual profit, and if everybody assumes everybody else to be exclusively interested in maximizing individual profit, i.e. with common knowledge of standard preferences, the tragedy of the commons materializes (Hardin 1968). Yet, happily, observations from the field and the lab are considerably less gloomy (Ledyard 1995, Zelmer 2003, Chaudhuri 2011). One reason is motivational. Experimental evidence shows that many participants are willing to forgo the opportunity to exploit their counterparts if they are sufficiently confident that they themselves will not be exploited. This is of course the definition of a conditional cooperator (Fischbacher, Gächter et al. 2001, Fischbacher and Gächter 2010).

Moreover, most situations that can be modelled as a social dilemma involve repeated interaction. As is well known, repetition is immaterial if it is known at which point the game ends and common knowledge of standard preferences is assumed. In the final round, every

participant will defect. This is anticipated by all other participants. Therefore, in the penultimate round, cooperation does not pay. Through unravelling, this holds for all earlier periods. Yet even if individually all participants are indeed exclusively interested in maximizing profit, there is scope for cooperation. It exists if sufficiently many participants are individually uncertain whether the remaining participants also hold standard preferences, and are willing to act upon them. If not, the famous gang-of-four result applies (Kreps, Milgrom et al. 1982). The folk theorem is applicable (Fudenberg and Maskin 1986, Fudenberg, Levine et al. 1994). Any level of contributions to the public good is an equilibrium. The original cooperation problem turns into a coordination problem. Participants must find a way of zeroing in on a plausible contribution level. Repetition makes cooperation even more likely if at least some participants are not only farsighted free riders, but are genuinely (conditionally) cooperative (Engel and Rockenbach 2024). Hence, with repetition, there is even more scope for socially beneficial type heterogeneity.

In this experiment, we start from repetition, and we allow for preference heterogeneity. But we add a complication that appears in many, if not most, real-life dilemma situations. The situation is dynamic, in the sense that the social problem exacerbates over time. We manipulate the speed at which the social problem becomes patent. We implement a public bad, not a public good, because doing so makes the design more intuitive (Sonnemans, Schram et al. 1998, Moxnes and Van der Heijden 2003, Engel 2015a). Of course, this is ultimately only a (valence) frame (cf. Dufwenberg, Gächter et al. 2011). We could have represented the same incentive structure as a public good that, over time, becomes harder and harder to provide.

In real life, many public bads have a dynamic component. More specifically, the frequency and intensity of harmful behavior in the past affects the amount of damages realized in the present. Take the quintessential public bad, damage to the environment. To a remarkable degree, the environment is forgiving. Nature recovers from occasional damaging events. Yet if damage becomes more frequent, nature gradually loses resilience and takes longer and longer to recover (Nelson, Adger et al. 2007, Moghim and Garna 2019). As another illustration, take the rules of conduct in a community. Most community members are willing to tolerate occasional transgressions and continue to follow the rules. Yet, if transgressions accumulate, the willingness of more rule-abiding members to keep up the group spirit by following the rules themselves will gradually decay. Community rules may cease to guide behavior (cf. the evidence on the so-called broken windows effect, e.g. Keizer, Lindenberg et al. 2008, Beckenkamp, Engel et al. 2009). Or consider a group that has established beneficial exchange with a second group, based on a mutual guarantee of exclusivity. If the groups are not too small, or if behavior is not too transparent, this beneficial arrangement may survive occasional outside dealing by single group members. Yet if such violations of the arrangement become frequent, the willingness of the counterpart group to reciprocate may deteriorate. In the long run, the mutual benefit may be lost and everybody will be worse off (if they succeed, this of course constitutes a cartel. For experimental evidence see Engel 2015b). In all of these scenarios, harm is not tangible immediately but takes time to realize. A single misdeed may, in and of itself, not result in noticeable detriment for society. But over the course of time and together with the transgressions of others, seemingly small acts can develop to have bigger consequences than initially observed. This deferred nature of harm realization alone may already affect the actors' ability to cooperate.

A final illustration of this is particularly timely: collective harm may also result from revealing personal information on the individual level. In the digital realm, sharing personal data can be construed as the contribution to a public bad (Fairfield and Engel 2015). Individual information feeds into pattern recognition algorithms; this enhances their ability to predict behavior and preferences, even for individuals who have limited their personal data disclosure. The improvement of these algorithms, while providing valuable tools, may also result in detrimental societal outcomes: they can be used to personalize prices (Dubé and Misra 2023) or to fine-tune choice architecture to manipulate online decision-making (Susser, Roessler et al. 2019). As the volume of disclosed information grows and predictive algorithms become more precise, the potential for exploitation grows, and collective harm may surpass individual benefits. This concern is magnified by the current rise of generative AI, which leverages vast datasets to produce highly accurate and individualized predictions, thus amplifying risks of misuse and raising profound ethical questions about privacy and consent.

**Design.** With field data it would be very difficult to investigate whether the dynamic nature of a public good, and the resulting delayed visibility of social damage, exacerbate a public bad. It would already be difficult to find otherwise identical public bads that only differ in the dynamics of the production function. And for identification one would, additionally, need pseudo-random assignment to either of these settings. The lab is attractive since the production function can be made the treatment manipulation, and participants can be randomly assigned to treatment.

The stage game in our experiment is a public bads game with fixed groups of four. Subjects receive an initial lump-sum endowment and can decide to contribute up to 15 “points” per round. Each contributed point creates an individual benefit, but also causes harm to every member of the group. If all participants contribute an equal amount, this harm outweighs the sum of individual benefits. The game is repeated 20 announced times.

In our experiment, however, the harm created by contributions does not fully materialize immediately. Instead of deducting the created damages promptly at the end of each round, it takes several periods for the damage to fully unfold. We manipulate the number of periods required for this process by altering the weight with which previous contributions are transferred to the next round. In our two treatments, this “decay factor”  $\delta$  is either .1 or .9 so that either 10% or 90% of the current period’s contributions are considered for deductions in the subsequent round. Over several periods, this mechanism creates a “stock” of contributions, which decays over time (with the speed of decay depending on the treatment). Because of this, contributions do not only cause harm in the present but also in later rounds. If participants continuously choose to contribute, this stock will grow over time, causing even higher deductions in each period. If participants stop contributing, the stock and the associated damages decrease over time.

We keep the economic consequences of contributing points constant between treatments. To achieve this, the introduction of another variable, the “multiplier”  $\mu$ , is required. This multiplier determines how much of the remaining stock is deducted from the respective periodic payoff of each group member. We adjust this multiplier to ensure that the harm for group welfare caused by each contributed point is the same irrespective of treatments

Therefore, fully prevoiant participants should react the same way to both treatments. The experiment is however motivated by the expectation that it becomes harder for groups to mitigate the dilemma if it takes longer for the social damage to materialize.

Additionally, our experiment is designed such that the ultimate economic consequences of contributions are the same in every period. Our experiment has a defined end. Without an adjustment, a contribution to the public bad that occurs later in the game would matter less for total harm. Moreover, with greater delay, the shadow of the future is longer. Consequently, without an adjustment in the treatment with greater delay, a bigger part of the theoretical damage would only materialize after the experiment has been closed. Observations from the two treatments would not be comparable. To resolve this issue, we introduce an “endgame deduction”: After the last period, the harm which the remaining stock would have caused, had the game continued indefinitely, is subtracted in a one-time payment from each group member’s profit. Through this payment, any unit contributed at any time and in either treatment eventually has the same detrimental effect on every group member. The endgame deduction thus keeps the long-term effect of contributions constant across treatments and over time.

As the last period is announced, we do not have to worry about multiple equilibria. The Folk Theorem does not apply. If we assume common knowledge of rationality, in the final round, every participant contributes fully. This is anticipated, which is why it does not pay to invest in cooperation in the penultimate round. By unraveling, the same reasoning holds for the second to last period. In the logic of backwards induction, all participants contribute fully right from the start.

While the main motive for the endgame deduction is experimental, plausible analogues for this mechanism exist in reality. Take the quintessential example of climate change: Even if carbon emissions are ceased, the pollution caused up to that point continues to have a negative impact on the environment, which society must endure. This collective harm only dissipates slowly, as nature regenerates over time. An alternative example are pattern recognition algorithms. An increase in their capabilities might result in harmful business conduct, but might lose relevance over time, as their predictive power slowly dissipates when not continuously fed with up-to-date information.

We inform participants about these mechanisms and do everything to make sure that they understand the incentive structure. We show them their individual payoff function (see Section 3) and represent the function with an intuitive example and a graphical description of how harm develops over time. Finally, we test participants’ understanding with an extensive set of instructive control questions.<sup>1</sup> Still, the dynamic nature of harm development arguably creates a cognitive challenge. This potentially adds on to the motivational challenge resulting from social dilemma.

To isolate the cognitive challenge, we introduce a second treatment dimension. Besides the *group* condition described above, we implement a *single* treatment. In this treatment, we replicate the otherwise identical experiment but replace human group members with three

---

<sup>1</sup> For more detail, please see Section 4 and the instructions in Appendix B.

“computerized players”. Participants are informed that, from the second period on, these automated players exactly copy the participant’s contribution choice from the previous period. If a participant has fully grasped the implications of the design, they should not contribute anything in every period but the last. For if all group members contribute the same amount, their individual payoff is lower than when contributing zero.

**Preview of results.** We find that, even in the *single* condition, multiple participants make positive contributions. Despite completely removing strategic uncertainty and making social preferences pointless, the fact that social harm is not immediately visible creates a problem. The dynamic nature of the game alone creates a cognitive challenge.

Descriptively, in the *group* condition, contributions to the public bad are higher with  $\delta = .9$  in all but a single period. Yet statistically, the level effect of treatment on contribution is not significantly different from 0. We do however find that, with  $\delta = .9$ , contributions increase more rapidly over time as compared to  $\delta = .1$ . Contrasting the *single* and *group* treatments, we also find a clear level effect on contribution: if the cognitive challenge is compounded by a motivational challenge, the social dilemma looms larger.

We find a clear effect of social preferences on contribution. Participants contribute less to the public bad if a subsequent test for social value orientation characterizes them as prosocial. Yet this effect does not differ between the  $\delta = .1$  and the  $\delta = .9$  condition.

Most importantly for our research question, the size of contributions is explained by personal experiences. In the four periods after a participant has received a negative period payoff for the first time, she significantly reduces her contributions – the more so the closer to this experience. This shows that participants must themselves *experience* that high contributions inflict harm on society, including themselves. Most interestingly, the treatment effect of  $\delta$  becomes insignificant once we add dummies to the statistical model for every period after the individual period payoff has first turned negative. This also holds when interacting the period dummies with  $\delta$ . Consequently, the decay factor does not matter per se. It matters because a high decay factor shields participants for a long time from bad individual experiences. Only once participants make these experiences do they reduce their contributions to the public bad.

The remainder of the paper is organized as follows: in the next Section, we relate our experiment to the literature. In Section 3, we introduce the design of the experiment. In Section 4, we present our hypotheses. In Section 5, we report results. Section 6 concludes with a discussion.

## 2. Related Literature

Two literatures in behavioral economics are relevant: the one on public bads, and the one on dynamic public goods. Since we construct our problem as a public bad, Andreoni (1995) provides a good starting point. He shows that reframing a static linear public good as a public bad substantially reduces cooperation, despite the fact that incentives are unchanged. Later

work replicates the effect (Sonnemans, Schram et al. 1998, Moxnes and Van der Heijden 2003, Khadjavi and Lange 2015).

There is also a developed literature studying the effect of varying the design of linear public goods games, including the number of group members (Isaac, Walker et al. 1994), the marginal per capita rate and hence profit from cooperation (Isaac and Walker 1988), the endowment (Laury, Walker et al. 1999), and repetition (Isaac, Walker et al. 1984). If the horizon of the future is longer, cooperation is maintained for a longer time (Gächter, Renner et al. 2008).

While there is a future, in such games each round gives participants a fresh start. As a matter of theory, Fershtman and Nitzan (1991), Wirl (1996), Kossioris, Plexousakis et al. (2008) note that the voluntary provision of public goods may be profitably studied dynamically. Cadigan, Wayland et al. (2011) test a very moderate version of a dynamic game. In their public good, the benefit from contributions to the public good ‘carries over’ to the next round. They find that this partly mitigates the dilemma. Duffy, Ochs et al. (2008) compare a static public good with one where the public benefit is allowed to accumulate over time. In the long run, groups do better in the dynamic setting, but they contribute less in the beginning. Battaglini, Nunnari et al. (2016) allow a public good or a public bad to accumulate over time, with no decay at all. They find under-provision of the public good, compared with the theoretical prediction.<sup>2</sup> Noussair and Soo (2008) investigate contribution behaviour in a public good game where previous cooperation affects marginal per capita rate in the current round. In their experiment, most groups do not show the typical declining trend in contributions. Finally there is a literature on threshold public bads. In these games, once the sum of contributions reaches the threshold, all is lost (Croson and Marks 2000, Bosetti, Heugues et al. 2017, Guilfoos, Miao et al. 2019).

As the cited theoretical literature shows, dynamic games often exhibit a multiplicity of equilibria, since the growth of the stock over time may change players’ best responses. Such approaches differ from ours in that we seek the simplest possible experimental test of how the decay factor of a public bad affects group coordination. In our experiment, it is always individually better to contribute maximally; the individually optimal choice of each player does not change as a function of the group history. The accumulation and decay factors merely change the speed at which the social dilemma unfolds its consequences, and thus the speed at which the public harm responds to coordination attempts. We are interested in the behavioral effects of these manipulations.

Studies of dynamic public bads sometimes focus on the sustainment of long-run cooperation in infinite games by providing a probability of continuation. Dal Bó and Fréchette (2011) vary this continuation probability, but the decision in each stage game does not directly affect the payoff in later stage games. Hence the realization of social harm is immediate, and not deferred. This is different in our experiment: while the number of rounds is fixed, contributions to the public bad accumulate over time, and are only partly experienced at the

---

<sup>2</sup> Although Diev and Hichri (2008) have the word “dynamic public good” in the title of their paper, they study a very different form of dynamics. Within each round of a static game, participants are allowed to adjust their contribution choices upwards, in the light of information about other participants’ (initial) choices.



point in time when contributions are made. We vary the rate with which the stock of social harm decays over time.

Public bad games with dynamics that somewhat resemble our experimental design have been conducted in the context of carbon emissions. Pevnitskaya and Ryvkin (2013), Pevnitskaya and Ryvkin (2022) match participants into groups of two and let them choose a “level of production”. Production grants individual revenue but creates “emissions” that induce social harm. These emissions are transferred to the subsequent round with a weight of 0.75, providing a mechanism akin to the decay factor  $\delta$  in our own experiment. However, their study focuses on the effect of framing and differing time horizons and does not manipulate the speed at which the social harm materializes. Ghidoni, Calzolari et al. (2017) conduct a similar experiment in groups of four. While they do induce public harm statically (i.e., realizing within a single round), this realization takes place either immediately or with a delay of two periods. They observe no significant effect on cooperation on balance, but do find a significant positive time trend for contributions when damages are delayed.

To the best of our knowledge, the study by Calzolari, Casari et al. (2018) presents the only other experiment thus far to compare different decay factors as treatments in dynamic games. Their experiment is motivated by a set of stylized facts about climate change: states discount future earnings, and investments into abatement have diminishing marginal returns (which is why they induce a discount factor of 0.92); states have heterogeneous preferences (which is why they give group members different profit functions); zero emissions are not a plausible goal (which is why they enforce a minimum level of emissions, and implement a socially optimal amount of individual emissions); in their “persistent” design, emissions today yield constant damage in every later period. This design leads to a complicated set of possible Markov-chain equilibria. They have a surprising finding: with maximum delay (“persistent”), first round emissions are significantly lower (which is the opposite of our expectation). In comparison, our experiment is (deliberately) much simpler. The dynamic frame notwithstanding, it is a standard linear public bad. Socially optimal contributions are 0. We have an announced number of rounds, so that the implied discount factor is 1 (or the implied discount rate is 0). All group members have the same profit function. All we manipulate are the experiences participants are making, and the moment in time at which the actual harm materializes. With this design we can isolate what we believe to be a critical difference between static and dynamic social dilemmas: whether individuals perceive the harm they inflict on the community while they do so.

Ultimately, it should be emphasized that the effect of deferring the materialization of harm we examine in this experiment is not a matter of time preferences. Previous research has pointed out that discounting tends to be hyperbolic rather than exponential (Laibson 1997), thus revealing a time-inconsistent preference structure of individuals. This tendency to weigh immediate rewards higher than future consequences (O’Donoghue and Rabin 1999) (O’Donoghue and Rabin 1999) has been described as a “present bias” (O’Donoghue and Rabin 2015). In empirical studies investigating this effect (see the extensive overview by Frederick, Loewenstein et al. 2002), participants are typically confronted with the choice of either receiving a smaller amount of money immediately or a larger amount at a later point in time. That is not the case in our experiment. All participants receive their payoff at the end of the respective session. Whether they lose their endowment in an earlier or later period of the

game is not a matter of time preferences and does not affect the duration participants have to spend in the lab. We merely manipulate the speed with which the collective harm materializes within the experiment.

### 3. Design

In our experiment, we employ a two-by-two factorial between-subjects design: As a first treatment dimension, we vary the decay factor between a *low* decay factor treatment ( $\delta = 0.1$ ) and a *high* decay factor treatment ( $\delta = 0.9$ ). For both treatments, we also manipulate whether subjects interact with each other in fixed groups of four (*group* treatment) or with a set of computerized group-members that copy the subject's contribution from the previous period (*single* treatment). In all treatments, participants receive an initial lump-sum endowment of 28.00 EUR and can make contribution choices for 20 announced periods.

**Group treatment.** In each period, participants in the group ( $N = 4$ ) can decide on a contribution  $c$  between 0 and 15 points,  $c \in \{0, 1, 2, \dots, 15\}$ . Each contributed point incurs an individual cost of 0.01 EUR but creates an individual benefit of 0.10 EUR ( $\rho = 10$ ). For any point invested by any group member, collective harm of 0.03 EUR occurs for the contributor and *every group member* ( $\hat{\mu} = 3$ ). Hence, regardless of treatment, the long-term consequences of contributing one unit  $c$  are always given by

$$\pi_i = (\rho - 1)c - \hat{\mu} \sum_{j=1}^N c_j$$

where  $\pi$  is period profit,  $\rho$  is the factor by which the contribution of 1 unit  $c$  pays back for the contributor  $i$ ,  $N$  is the number of group members and  $\hat{\mu}$  is the long-term detriment from investing one unit, which is held constant across treatments.

The individual costs and benefits from contributing are immediately realized after each period  $t$ . The collective harm is not. Instead, the materialization of the collective damages is partly deferred to later rounds. For this, we manipulate the rate  $\delta$  at which contributions made in period  $t$  lose their impact on the harm participants experience in later periods  $t + T$ . The decay factor  $\delta$  therefore describes the weight with which the stock of contributions is transferred to the subsequent period. To keep the long-term consequences of contributions consistent between treatments, we introduce a multiplier  $\mu$  which determines how much of the stock of contributed points is deducted from each group member's period profit. Therefore, in each period payoff is defined by

$$\pi_{i,t} = (\rho - 1)c_{i,t} - \mu(C_t + \delta C_{t-1} + \delta^2 C_{t-2} \dots + \delta^T C_{t-T})$$

where  $C_t$  is the sum of contributions all group members have made in period  $t$ , and  $T$  is the sum of periods the game has lasted.

We adjust  $\mu$  such that, irrespective of treatment, we have

$$\hat{\mu} = \frac{\mu}{1 - \delta} = 3$$

from the definition of the sum of an infinite geometric series. We guarantee this geometric series to be correct by introducing an “endgame deduction”. Because the materialization of the collective harm is delayed, it is unlikely to fully realize within the announced duration of 20 periods – especially in the *high* decay treatment ( $\delta = .9$ ), in which damages are more deferred. With the one-time payment, we deduct the stock of remaining contributions at the end of the game from each group member’s profit. This allows us to simulate an indefinite continuation of the negative consequences induced by prior contribution behavior, even beyond period 20. This deduction ensures that each point contributed throughout the 20 periods does indeed cause the same collective harm of 0.03 EUR. It allows us to keep the economic consequences of contributions constant across treatments and over time.

Thus, while keeping the overall harm of  $\hat{\mu} = 3$  constant, we change the time required for the collective harm to fully materialize. Table 1 below provides an overview of the differences in these parameters between treatments.

Treatment	$\delta$ (decay factor)	$\mu$ (period multiplier)	$\hat{\mu}$ (long-term multiplier)
“low” decay	0.1	2.7	$\frac{\mu}{1 - \delta} = 3$
“high” decay	0.9	0.3	

Table 1  
Difference of Decay Factors Across Treatments

Figure 1 shows the dramatic consequences that the different treatments have for the participants’ experiences. With  $\delta = .1$ , almost all individual and social harm is felt immediately. By contrast with  $\delta = .9$ , harmful acts do not have serious consequences for a very long time.

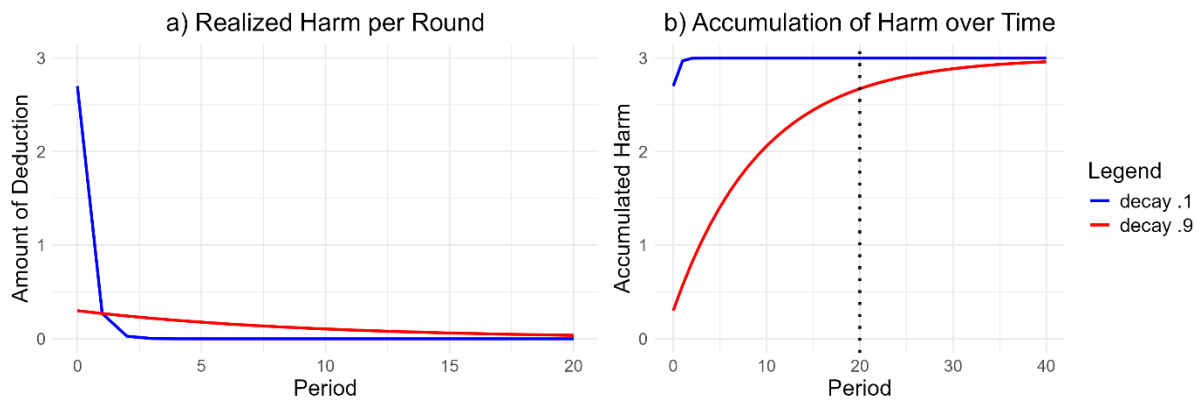


Figure 1  
Collective Harm for One Point Contributed  
left panel: Harm per Round  
right panel: Accumulated Harm

After each period, participants receive feedback on their individual contribution, the other members' aggregated contributions, the total group contribution, their individual period revenue (without deducting collective harm) and their individual period earnings (including the deduction of collective harm), before deciding on contribution for the subsequent period.

**Single treatment.** In our *single* treatment, participants are not randomly assigned to groups of four human participants. Instead, they interact with "computerized group members". Human participants are informed that the three automated players always copy the respective participant's contribution behavior from the previous period. E.g., should a participant invest 7 points in round 2, the remainder of the group will invest 21 points in round 3. The automated group members' contribution is randomized in the first period, as there is no behavior to be copied. We chose a random initial computer contribution over a maximum or minimum contribution to not induce any normative benchmarks. We run this treatment dimension to isolate the cognitive dimension of our experiment from the motivational one: while the cognitive challenge of grasping the long-term effects of contributions remain comparable when interacting with a computer, the motivational challenge of anticipating group and coordination behavior is dropped, as the computer perfectly copies previous behavior.

**Information and comprehension.** We inform participants of all the abovementioned mechanisms before starting the experiment: their individual periodical payoff function, the number of periods to play, whether they are paired with other participants or a computer, and, in case of the latter, how the computer will behave. We also explain the dynamic development of harm over time, and the consequences this entails for themselves and their group members, including the endgame deduction. We explicitly highlight  $\hat{\mu}$  by stressing that, because of the dynamic development of harm, every invested point leads to a deduction of 0.03 EUR for every group member in the long run, including themselves.

While we are fully transparent, the experiment arguably remains complex for many participants. We therefore pay particular attention to making sure that every participant fully understands the mechanism at hand and thus, the consequences of their actions. To do so, we make the experimental instructions as accessible as possible, by including several written examples of the design's key elements and by providing a graphical exemplification of collective harm development over several rounds. Moreover, we implement a set of extensive (and arguably challenging) control questions that require serious understanding of the mechanisms at hand. These questions are programmed in an instructive manner: they require the step-by-step calculation of individual period payoff and the development of collective harm over several periods, in order to break down the design into pieces that are easier to understand. The control questions require the calculation of the subject's individual payoff over time as well as the long-term consequences of contributions, to make sure every participant is aware of this mechanism. The long-term effects of contributions and the existence of the endgame deduction are particularly stressed in one control question, which asks for the long-term group consequences of investing 12 points in the beginning and towards the end of the game – the response to both was a deduction of 36 points. In the *single* treatment, we add an additional question about the contribution behavior of automated group members.

Whenever participants provided an incorrect answer, a hint was displayed which pointed towards the part of the design relevant for the respective question. This was done to improve participants' comprehension of the mechanism without giving away the correct answer.<sup>3</sup> Additionally, this helped to reduce the requirement for personal involvement of the experimenter and thus ensured a comparable level of support for all participants in need of further clarification. The decision phase only started once every participant in the session had successfully responded to all the control questions. The time required from handing out the instructions until the beginning the experiment was 31 minutes on average per session.<sup>4</sup> Considering the average session duration of 83 minutes, this demonstrates the considerable focus we placed on making sure that every participant is aware of the consequences of their choices.

**Further data collection.** After conducting the dynamic public bad game described above, we elicited the subjects' second-order beliefs regarding the social acceptability of differing levels of contribution behavior (Krupka and Weber 2009) and their social value orientation using the 6-item SVO slider measure (Murphy, Ackermann et al. 2011). In a subsequent survey, we gathered subjects' personality traits using the BFI-10-Item-Scale (Rammstedt and John 2007), political orientation using the POLID-Scale (Ulrich 2021) and concerns regarding online privacy (Dienlin and Trepte 2015). We also elicited demographic information on the subjects' age, gender, siblings, previous experimental experience, student status and field of study as well as previous work experience. The mean age in our panel was 26.26 years; 81.13% of participants were students; 53.17% identified as female and 5.16% identified as diverse or chose not to disclose their gender. For more information, please see Tables A1 and A2 in Appendix A.

Data collection took place in 12 sessions with an average of 21 participants between the 25.06.2024 and the 04.07.2024 at the Decision Lab located at the Max Planck Institute for Research on Collective Goods in Bonn, Germany. Of the 252 participants, 60 were assigned to the *single* treatment and 192 were assigned to the *group* treatment, with a balanced split between the two decay factors in both the *single* (30/30) and the *group* treatment (96/96). The average payment for participation was 27.40 EUR (29.32 EUR in *single* treatment and 27.02 EUR in *group* treatment). The experiment was conducted on a computer, using the zTree software (Fischbacher 2007). Participants were invited via hroot (Bock, Baetge et al. 2014).

#### 4. Hypotheses

**Theoretical starting point.** If we assume common knowledge of standard preferences, participants only care about their own profit; they assume all other group members to do the same; they fully anticipate the long-term consequences of their own choices; they fully anticipate the long-term consequences of the choices they expect other participants to make.

---

<sup>3</sup> Please refer to the experimental instructions and control questions in Appendix B. Screenshots of the experimental stages, including the control questions, can be found in the online materials, see [https://osf.io/hqawx/?view\\_only=709a46333db6400086393ee1f5844f59](https://osf.io/hqawx/?view_only=709a46333db6400086393ee1f5844f59).

<sup>4</sup> Please note that this measures the slowest participant in each session. The individual average to clear the control questions was 14.61 min with a standard deviation of 5.35 min.

The design holds the long-term consequences of choices constant across treatments. In our experiment

$$\frac{\rho - 1}{N} < \hat{\mu} < \rho - 1$$

defines the social dilemma. This gives us our null hypothesis which, through unravelling, also holds in the repeated game:

**$H_{0_{group}}$ : common knowledge of standard preferences:** Participants invest the maximum amount in all periods. There are no treatment differences.

In the *single* condition, through their own contribution in period  $t - 1$ , participants control the contributions of the remaining group members in period  $t$ . This effectively removes the social dilemma. Like a social planner, participants can maximize their own share of group profit

$$\pi_i = \frac{1}{N} \left[ N(\rho - 1)c_i - N\hat{\mu} \sum_{j=1}^N c_j \right]$$

For optimal group profit, social harm matters. Hence total group contributions count  $N$  times. Given  $N\hat{\mu} > (\rho - 1)$ , the human participant will refrain from investing, except for the final period, in which there is no shadow of the future anymore:

**$H_{0_{single}}$ : control over group contributions:** If they can control how much others contribute, participants invest nothing in all rounds but the final; they invest fully in the final round. There are no treatment differences.

**Preregistered hypotheses.**<sup>5</sup> While a perfectly foresighted participant should only care about the long-term consequences, and hence about  $\hat{\mu}$ , we have run the experiment as we were expecting that participants would be influenced by the earning prospects in individual periods. While keeping  $\hat{\mu}$  constant, we manipulate the rate at which this harm materializes. We envisage that this difference affects the choices participants make. Arguably, anticipating the future effect of present contributions to the public bad is cognitively more demanding if they do not materialize immediately. It requires that participants, at least in a summary fashion, work themselves through the implications of choices in one period for their final profit. With  $\delta = .1$ ,  $\mu = 2.7$  most of individual and social harm resulting from investment materializes immediately, and it is severe. It is easy to see that even  $4 * 2.7 = 10.8 > 9$ . Hence the social dilemma hits home right from the start of the experiment. By contrast with  $\delta = .9$ ,  $\mu = .3$ , the immediate effect of every group member investing 1 unit is minimal:  $4 * .3 = 1.2 < 9$ . The social dilemma takes much longer to materialize. If the higher decay factor makes it harder to understand the implications of the design, we expect

---

<sup>5</sup> Hypotheses  $H_1$  throughout  $H_3$  were pre-registered in this exact wording, see [https://osf.io/hqawx/?view\\_only=709a46333db6400086393ee1f5844f59](https://osf.io/hqawx/?view_only=709a46333db6400086393ee1f5844f59).

**H<sub>1</sub>: effect of decay factor on level of contributions:** Individuals' contributions to the public bad are higher when the decay factor is higher

This effect should hold in the *single* and in the *group* condition, as this cognitive challenge is present in both conditions.

**H<sub>1</sub>** requires at least a rudimentary form of anticipation. **H<sub>2</sub>** is less ambitious. It expects that, actually, participants only react to experience. The higher  $\delta$  (and the smaller  $\mu$ ), the longer it takes until the long-term effects of social harm translate into a clear effect on per period payoff. We therefore predict

**H<sub>2</sub>: effect of decay factor on slope of contributions:** In the higher decay factor treatment, efforts to cooperate – if visible – will occur later and are less pronounced as compared to the low decay factor treatment.

As we have explained, with common knowledge of rationality, the predictions for the *single* and the *group* treatment differ radically: while participants should contribute nothing (except for the final period) if they are *single*, they should contribute fully if they are a member of a *group* of agents acting strategically. Actually, we expect only a fraction of participants to fully grasp the strategic implications of the *single* treatment. Still the effects of one's own contributions are easier to understand if one knows that the remaining group members will strictly copy last period's choice. It is also easier to learn from experiences. This is why we predict

**H<sub>3</sub>: single vs. group condition:** In the *group* treatment, individuals will contribute more than in the *single* treatment.

**Motivated reasoning.** The main reason for running the *single* treatment is isolating cognitive effects. Participants have no reason to care about the fictional payoff reaped by the three computer group members. By contrast, in the *group* treatments, cognitive and motivational effects are likely to compound. As we implement a social dilemma, if a participant holds social preferences, she should take into account the effects of her own choices on the payoff of the remaining group members (for a summary account, see Fehr and Schmidt 2006). Social preferences can be interpreted as sensitivity to fairness, either in payoffs (Fehr and Schmidt 1999, Bolton and Ockenfels 2000) or in intentions (Rabin 1993, Falk, Fehr et al. 2008). People tend to interpret fairness problems in a self-serving manner (Konow 2000, Konow 2003). Such a self-serving bias is made possible by perceived ambiguity (Sarin and Weber 1993, Etner, Jeleva et al. 2012). A participant preserves a positive self-image while, at the same time, maximizing profit, by convincing herself that she actually is not doing any relevant harm (Haisley and Weber 2010). The very fact that, with a high decay factor, the long-term effects of selfish behaviour on the payoff of the remaining group members remain concealed for a long time creates scope for such motivated reasoning (Kunda 1990, Epley and Gilovich 2016, Dieckmann, Gregory et al. 2017). This is why we expect additional support for **H<sub>1</sub>**: it becomes easier for a participant holding social preferences to convince herself that, actually, contributing fully is not so detrimental for others; for **H<sub>2</sub>**: it takes a longer time before a participant holding social preferences sees with her own eyes that the group is suffering; and

for  $H_3$ : in the *single* condition, there is no scope for social preferences in the first place, and therefore also not for an additional effect resulting from motivated reasoning.

## 5. Results

**Single.** We begin with the design that isolates the cognitive challenge, i.e. with the *single* condition. Descriptively and statistically, we see no effect of treatment (Table 2 Model 1), which we would have expected under  $H_1$ . However, the effect on slopes turns out significant (see the interaction effect in Table 2 Model 2). Hence, we support  $H_2$ .

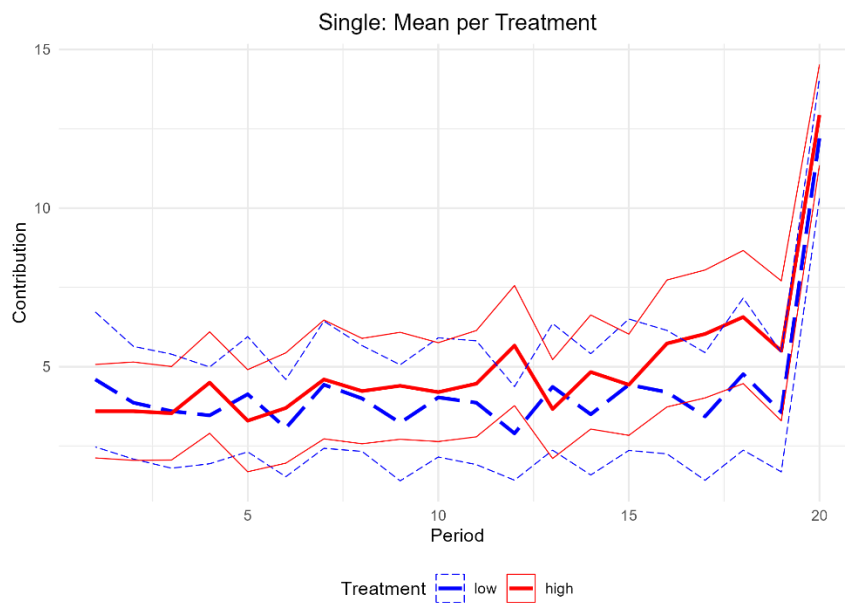


Figure 2  
Single Treatment: Aggregate Contributions with 95% CIs

	model 1 dummy 1 = normative choice	model 2 contribution (period 20 omitted)
high	-.153 (.106)	-.599 (1.020)
period		.005 (.025)
high * period		.129*** (.035)
cons	.508*** (.075)	3.821*** (.0721)
N uid	60	60
N obs	1200	1140

Table 2  
Hypothesis Tests for Single Treatment  
model 1: linear probability model with participant random effect



dv: dummy that is 1 for any period (from 1 to 19) if participant contributes 0, and is 1 if participant contributes the maximum in period 20  
model 2: linear model with participant random effect, period 20 omitted  
standard errors in parenthesis  
Hausman test insignificant  
\*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$

The individual contribution paths reported in Figure 3 are even more interesting. Recall that a participant who has thoroughly understood the design of the *single* experiment can easily maximize her income from the experiment: in period 1 she contributes 0 and keeps doing so until period 19. This guarantees that she keeps the maximum amount of her initial endowment. As she knows that the three computer players will copy her choice from period 19, in period 20 she contributes the maximum permissible amount of 15, which gives her an extra 0.90 EUR. Figure 3 shows that 10 participants in the *low* condition, and 5 participants in the *high* condition indeed behave in this way. A few more participants come close. But in both conditions, the majority behave in non-standard ways. Many of them even zig-zag, as if they had to learn the best response, although in the instructions share all the information they need to find it. Even in the absence of any strategic uncertainty, and in the absence of any motivational concerns, the mere dynamic character of the game makes it hard for the majority of participants to behave in the individually optimal way. The dynamic nature of the game creates a severe cognitive challenge.



Figure 3  
Single Treatment: Individual Contributions

**Group.** Descriptively, the average pattern of choices looks different when computer participants are replaced by human participants (compare Figure 2 with Figure 4).<sup>6</sup> Contributions generally show an increasing trend over time, with a notable incline in the beginning. For all but one period, average contributions in the *high* condition are above the contributions in the *low* condition. But this difference is small and insignificant (Table 3 model 1). Hence, we again do not support **H<sub>1</sub>**. But also in the *group* condition, contributions in the *high* condition increase more quickly over time than in the *low* condition. As the interaction effect in model 2 of Table 3 shows, the difference in slopes is significant. Thus, also for the *group* condition, we support **H<sub>2</sub>**.<sup>7</sup>

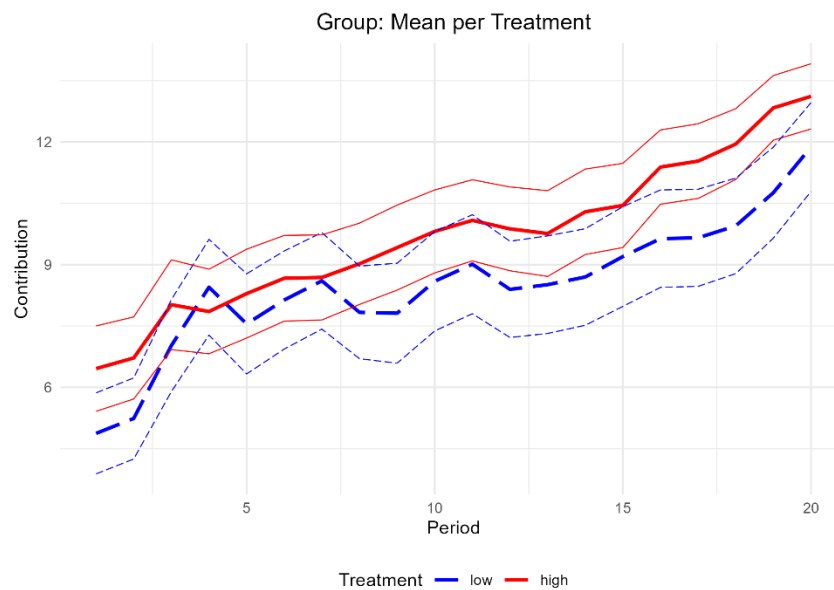


Figure 4  
Group Treatment: Aggregate Contributions

	model 1 contribution	model 2 contribution
high	1.221 (.922)	.600 (.947)
period	.274*** (.010)	.244*** (.015)
high * period		.059** (.021)
cons	5.613*** (.0661)	5.923*** (.675)
N subjects	192	192
N groups	48	48
N obs	3840	3840

<sup>6</sup> For a descriptive overview of average contributions per group, see Figure A1 in Appendix A.

<sup>7</sup> In the Appendix A3 we report an alternative specification with period and period<sup>2</sup>, interacting both with the *high* condition dummy. In this model, only the effect of period turns out significant. For a complete report including the pre-registered covariates, please see tables A4 – A6 in Appendix A.

Table 3  
Hypothesis Tests for Group Treatment  
Linear Mixed Effects Model  
Standard errors for choices nested in individuals nested in groups in parenthesis  
Hausman test insignificant  
\*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$

**Comparison of single with group condition.** Finally, Figure 5 shows that we have strong support for  $H_3$ : as expected, in the *group* condition, contributions are much higher, and they also increase much more quickly over time; statistical tests are in Appendix A (Table A7). If the dynamic game is compounded by strategic uncertainty, it is even harder to mitigate.

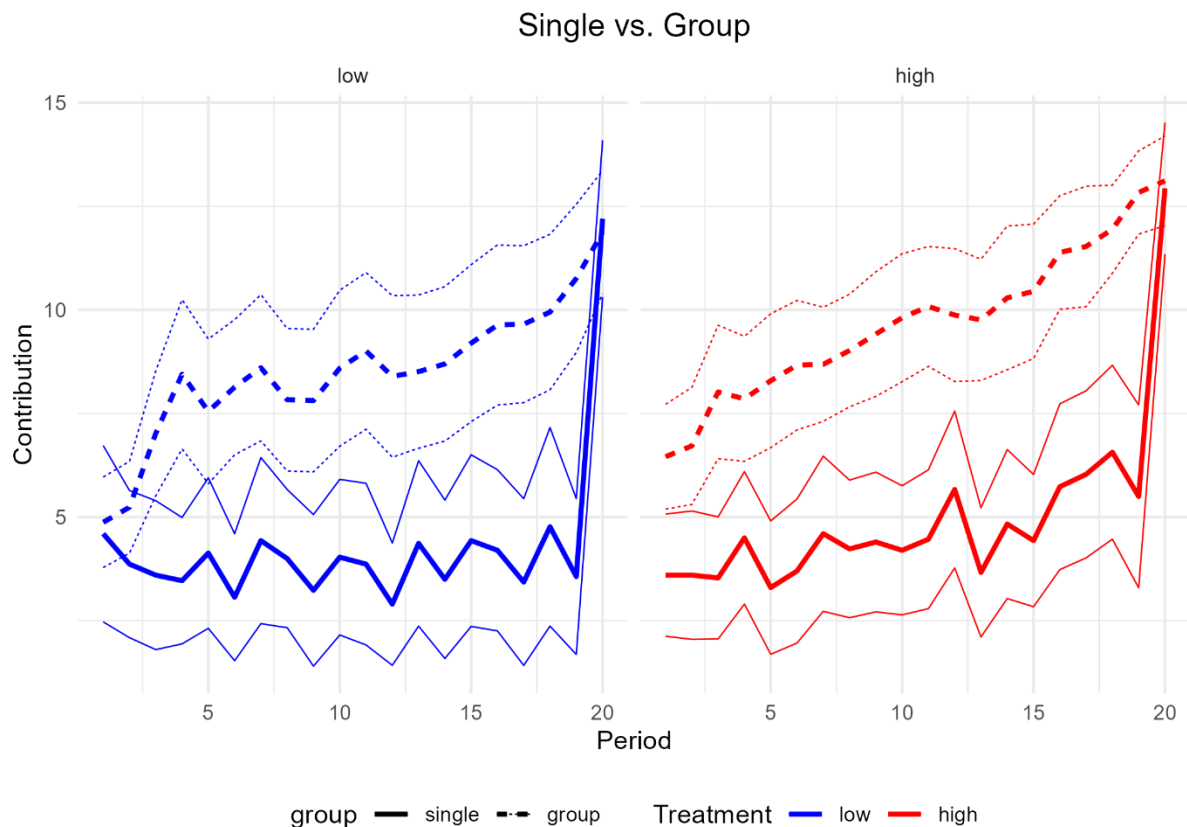


Figure 5  
Comparison of Single with Group Condition: Contributions

**Motivational types.** After the main experiment, we have tested participants on their social value orientation. As Figure A2 in Appendix A shows, we have a clear bifurcation: the majority of participants is prosocial (the mode of the distribution is at  $38^\circ$ ), while a sizeable minority is individualistic (the second mode is at  $8^\circ$ ). Using the canonical definition of types (with the cutoff at  $22.45^\circ$ ),<sup>8</sup> this implies that we have 133 prosocial, and 59 individualistic types in our data. Unsurprisingly, the individualistic participants contribute more. But we do not find a significant interaction between being individualistic and the decay factor (see Table A8 in Appendix A). This is a remarkable finding: the reason for the faster aggravation of the social

<sup>8</sup> For the translation of angles into types, see <http://ryanomurphy.com/styled-2/styled-4/index.html>.

problem in the *high* condition is not the presence of individualistic players. The decay of cooperation is not driven by preference heterogeneity.

**Complexity and comprehension.** The fact that individualistic social value orientation provides a strong and significant predictor for higher contributions in every model we test (see Table A8 in Appendix A) already indicates that participants, on average, know what they are doing. We additionally find that participants' comprehension of the design is no significant predictor for contributions in the group treatment. To test this, we consider the time participants needed to clear the control questions as a proxy for their individual degree of comprehension.<sup>9</sup> The data from our *single* treatment strongly suggests that this information does indeed capture individuals' comprehension: in the *single* treatment, one distinct contribution-pattern is always payoff maximizing, i.e., contributing nothing from period 1 to 19 and full contribution in period 20. Participants that fully understand the strategic implications of the experiment have no incentive to deviate from this pattern. Unsurprisingly, the time required to clear the control questions always constitutes a significant predictor for profit maximizing behavior in the *single* treatment, irrespective of the decay rate (see Table A9 in Appendix A).

However, we find no similar effect in our *group* treatment. When participants interact with one another in groups, the proxy for individual comprehension does not significantly predict contribution behavior, even when interacted with the social value orientation or the treatment variable (see Table A10 in Appendix A). This shows that the observed differences in cooperation cannot be explained with a lack of individual understanding.

**Effect of experience.** We support  $H_2$ , but not  $H_1$ : if it takes longer until the social harm fully materializes ( $\delta = .9$ ), the level of contributions is not significantly higher, but contributions increase faster over time. Our data suggests that the source of the problem is neither anticipation nor a lack of comprehension. Instead, as will be shown, we find that this difference can be explained by the absence of personal experience.

With  $\delta = .1$ , if all group members contribute fully all the time, individual period payoff is already negative in the first round.<sup>10</sup> By contrast, with  $\delta = .9$ , if all group members contribute fully all the time, individual period payoff stays positive until period 13.<sup>11</sup> Hence with  $\delta = .9$ , participants may nurture the illusion for a long time that, actually, not much harm is done. Not all participants in all groups contribute fully all the time. The point in time when individual period payoff for the first time turns negative exhibits variance. Yet as Figure 6 shows, the individual experience of negative period feedback has a striking effect: irrespective of treatment, participants respond immediately by drastically reducing their contribution.<sup>12</sup> In both conditions, it takes quite some time before contributions are back to the level observed before this negative experience.

---

<sup>9</sup> See Figures A3.1 and A3.2 in Appendix A for the distribution of the time needed to respond to the control questions in the *single* and the *group* treatment. N.B. that the *single* treatment included an additional control question regarding the automated players behavior.

<sup>10</sup>  $9 \times 15 - 2.7 \times 60 = -27$ .

<sup>11</sup>  $9 \times 15 - 0.3 \times \sum_{t=0}^{t=12} 0.9^t \times 60 = 0.75$

<sup>12</sup> 4 out of 96 participants in the .1-treatment and 14 out of 96 participants in the .9-treatment have either never experienced a negative period payoff or only did so in the ultimate period.

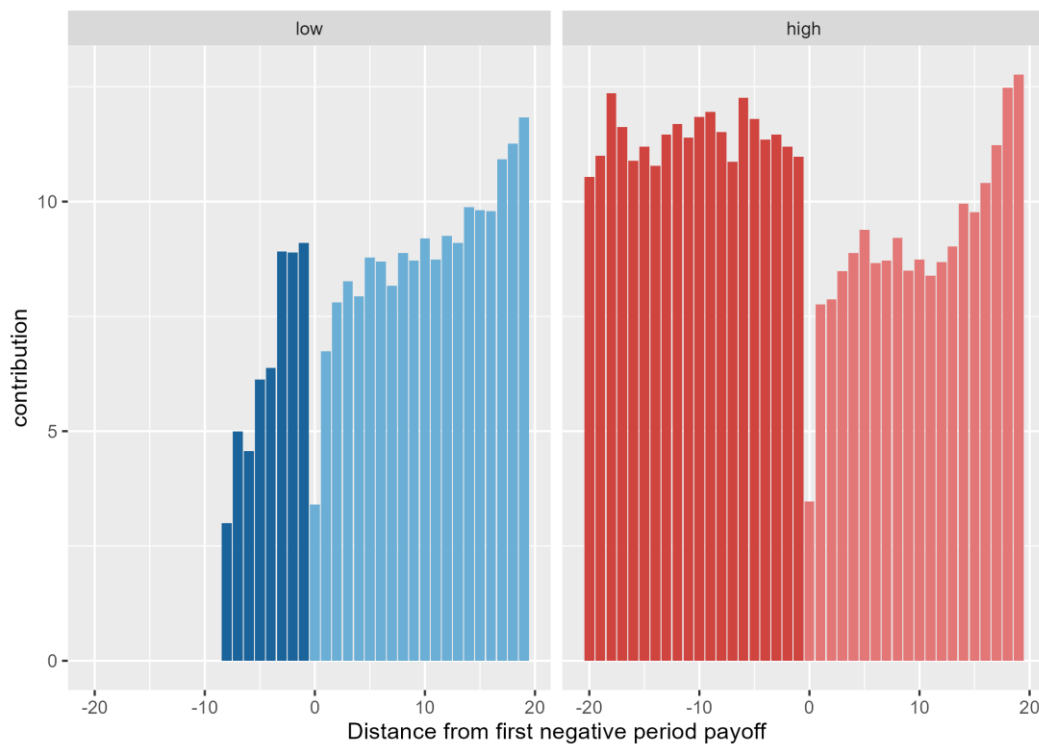


Figure 6

*Response of Contribution to Experience of Negative Period Payoff. Lighter colors indicate contribution behavior after experiencing a negative period payoff for the first time*

Table 4 provides statistical tests. Statistically, for the four periods following the first individual negative experience, contributions are significantly lower. Most interestingly, both the main effect of treatment *high* and all interactions are insignificant. This shows that the treatment effect is indeed driven by the apparent invisibility of social harm. The mechanism that has motivated the experiment is fully supported by the data. Notably, we do not find a similar contribution pattern in the *single* condition (see Table A11 in Appendix A), which indicates that this mechanism is mostly driven by the *motivational* challenges arising from the interaction with other participants.

	model 1	model 2
high	1.166 (.927)	.909 (.938)
1 period ago	-2.162*** (.318)	-2.615*** (.438)
2 periods ago	-1.547*** (.318)	-1.550*** (.438)
3 periods ago	-.970** (.320)	-1.093* (.438)
4 periods ago	-.938** (.322)	-1.419** (.438)
5 periods ago	-.223	-.571

	(.324)	(.438)
6 periods ago	-.559 (.327)	-.658 (.438)
7 periods ago	-.783* (.330)	-1.191** (.438)
8 periods ago	-.094 (.333)	-.474 (.438)
9 periods ago	-.398 (.337)	-.637 (.438)
10 periods ago	.082 (.342)	-.158 (.438)
high * 1 period ago		.950 (.637)
high * 2 periods ago		-.006 (.637)
high * 3 periods ago		.253 (.641)
high * 4 periods ago		1.036 (.646)
high * 5 periods ago		.755 (.650)
high * 6 periods ago		.205 (.658)
high * 7 periods ago		.921 (.665)
high * 8 periods ago		.871 (.673)
high * 9 periods ago		.547 (.685)
high * 10 periods ago		.566 (.702)
cons	8.854*** (.659)	8.987*** (.664)
N group	48	48
N uid	192	192
N obs	3840	3840

*Table 4*  
*Explaining Contributions with Having Experienced a Negative Period Payoff*  
 Linear Mixed Effects Model  
 Standard errors for choices nested in individuals nested in groups in parenthesis  
 Hausman test insignificant  
 \*\*\* p < .001, \*\* p < .01, \* p < .05

## 6. Discussion

Sometimes it helps to wait things out. At some future point in time, a problem may be easier to deal with. The person who has to face the problem may be in a better position then.

Circumstances may have developed favorably. But turning a blind eye to a problem may be, in retrospect, a bad idea. Facing the problem once it first becomes visible may be unpleasant. However, the hassle and the cost are often much greater if one has ignored the problem for a while.

Yet as long as the problem exclusively affects the individual, society, and the law for that matter, cannot easily claim a mandate for intervention. Procrastination may be unwise, but people are not ethically or legally obliged to be wise. This is different if individual neglect causes harm to other members of society. The ultimate reason for intervention is of course the externality. Externalities often result from the presence of a public good. Individual neglect multiplies in that every member of the relevant society suffers. However, the social problem aggravates when the effect of neglect is not immediately realized. If harm is gradual and delayed, individuals lack the immediate feedback needed to motivate corrective action, making the problem much harder to contain later on.

**Findings.** Our experiment isolates the nonlinear development of harm over time. We compare one treatment in which contributions to a public bad hit home quickly, with another treatment where individual and social detriment remains very mild for a substantial number of periods. We hold everything else constant so that we can see the pure effect of the production function of the public bad. Not only do we find the expected effect: contributions to the public bad are higher and more persistent when social harm takes time to materialize. Our results also pinpoint the mechanism driving this effect: participants only try to get the social problem under control once they have individually experienced that the public bad is real.

Our findings add nuance to the existing literature on dynamic public bads and public goods games. It extends prior work by uniquely isolating the decay factor's influence on participants' capability to coordinate, highlighting the role of delayed materialization of harm on cooperation. Our data show that it is a cause for concern if the true severity of a social problem is hidden or delayed. This suggests that interventions should focus on enhancing the immediacy of social harm – either by providing timely feedback or by internalizing the expected harm in advance to facilitate preventing an escalation of social dilemmas.

In many real-life scenarios, this conclusion is likely to be reinforced by the compounding effect of time preferences. The discounting of future values is empirically likely to be hyperbolic (Laibson 1997). This tendency towards weighting immediate rewards more heavily than future consequences (O'Donoghue and Rabin 1999) may exacerbate coordination failures, as participants prioritize short-term gains over mitigating long-term harm. These effects are likely to add on to the effects of delay shown here, making coordination attempts to address social harm even more challenging.

**Limitations.** While our experiment successfully isolates the impact of delayed harm in social dilemmas, it has certain limitations which should be acknowledged. First, the controlled lab setting, while ideal for examining specific causal mechanisms, cannot fully capture the complexity and variability of real-world social dilemmas. Our fixed 20-period setup, though controlled, does not fully capture potential long-term effects that might occur in an open-ended scenario in the field. In those cases, dynamic effects will usually be much more complicated. Additionally, individuals may receive mixed or misleading feedback about the

consequences of their actions, further weakening their willingness to act in a socially responsible manner. But as always, experiments are not meant to map reality. They are tools to isolate effects. In that spirit, with the design of our experiment we have ruled out alternative explanations as best we can.

A second limitation lies in the complexity of the experiment at hand. We have paid considerable attention to maximizing the participants' comprehension of the design by providing accessible instructions and implementing instructive but challenging control questions. Yet, the dynamic development of harm requires the consideration of two multipliers over time, which might be inherently challenging for some of the participants. While our data suggests that a lack of comprehension is not associated with individual contribution choices, it cannot be completely ruled out that some individuals may have not fully understood all implications of the design at hand. Yet, the challenges associated with complexity extend to all treatments in this experiment and may not explain the systematic differences in contribution behavior we have observed. Additionally, this reservation is not completely removed from reality: depending on the specific context, not all individuals may be fully aware of the precise extent to which their behaviour today will impact themselves and society in the long run.

Ultimately, as is usually the case with empirical research, our data does not only deliver answers, but also raises new questions. While the interaction effect of treatment and period is always positive and significant, indicating that the coordination is consistently more difficult over time in the .9-treatment, it is somewhat surprising that this effect is considerably larger in the *single* treatment (0.129\*\*\*, see Table 1 Model 2) as compared to the *group* treatment (0.059\*\*, see Table 2 Model 2). If the deferred realization of harm has a negative effect on the collective ability to overcome social dilemmas, one would expect this effect to be larger when the cognitive challenge of dynamic harm realization is compounded by the motivational challenge of anticipating coordination behavior, i.e., in the *group* condition. We leave further exploration of this finding to future work. Future research could also build on to these insights by exploring different forms of public bads in framed settings, as well as varying feedback mechanisms, to determine the policy interventions most suitable to enhance coordination in the context of deferred social dilemmas.



## References

- Andreoni, James (1995). "Cooperation in Public-Goods Experiments: Kindness or Confusion?" American Economic Review **85**: 891-904.
- Battaglini, Marco, Salvatore Nunnari and Thomas R Palfrey (2016). "The Dynamic Free Rider Problem: A Laboratory Study." American Economic Journal: Microeconomics **8**(4): 268-308.
- Beckenkamp, Martin, Christoph Engel, Andreas Glöckner, Bernd Irlenbusch, Heike Hennig-Schmidt, Sebastian Kube, Michael Kurschilgen, Alexander Morell, Andreas Nicklisch, Hans-Theo Normann and Emanuel Vahid Towfigh (2009). First Impressions Are More Important Than Early Intervention. Qualifying Broken Windows Theory in the Lab. Bonn, Max Planck Institute for Research on Collective Goods.
- Bock, Olaf, Ingmar Baetge and Andreas Nicklisch (2014). "Hroot: Hamburg Registration and Organization Online Tool." European Economic Review **71**: 117-120.
- Bolton, Gary E. and Axel Ockenfels (2000). "Erc: A Theory of Equity, Reciprocity and Competition." American Economic Review **90**: 166-193.
- Bosetti, Valentina, Melanie Heugues and Alessandro Tavoni (2017). "Luring Others into Climate Action. Coalition Formation Games with Threshold and Spillover Effects." Oxford Economic Papers **69**(2): 410-431.
- Cadigan, John, Patrick T. Wayland, Pamela M. Schmitt and Kurtis J. Swope (2011). "An Experimental Dynamic Public Goods Game with Carryover." Journal of Economic Behavior & Organization **80**: 523-531.
- Calzolari, Giacomo, Marco Casari and Riccardo Ghidoni (2018). "Carbon Is Forever: A Climate Change Experiment on Cooperation." Journal of Environmental Economics and Management **92**: 169-184.
- Chaudhuri, Ananish (2011). "Sustaining Cooperation in Laboratory Public Goods Experiments. A Selective Survey of the Literature." Experimental Economics **14**(1): 47-83.
- Croson, Rachel TA and Melanie Beth Marks (2000). "Step Returns in Threshold Public Goods. A Meta- and Experimental Analysis." Experimental Economics **2**: 239-259.
- Dal Bó, Pedro and Guillaume R. Fréchet (2011). "The Evolution of Cooperation in Infinitely Repeated Games. Experimental Evidence." American Economic Review **101**: 411-429.
- Dieckmann, Nathan F, Robin Gregory, Ellen Peters and Robert Hartman (2017). "Seeing What You Want to See. How Imprecise Uncertainty Ranges Enhance Motivated Reasoning." Risk Analysis **37**(3): 471-486.
- Dienlin, Tobias and Sabine Trepte (2015). "Is the Privacy Paradox a Relic of the Past? An in-Depth Analysis of Privacy Attitudes and Privacy Behaviors." European Journal of Social Psychology **45**(3): 285-297.
- Diev, Pavel and Walid Hichri (2008). "Dynamic Voluntary Contributions to a Discrete Public Good. Experimental Evidence." Economics Bulletin **3**: 1-11.
- Dubé, Jean-Pierre and Sanjog Misra (2023). "Personalized Pricing and Consumer Welfare." Journal of Political Economy **131**(1): 131-189.
- Duffy, John, Jack Ochs and Lise Vesterlund (2008). "Giving Little by Little. Dynamic Voluntary Contribution Games." Journal of Public Economics **91**: 1708-1730.
- Dufwenberg, Martin, Simon Gächter and Heike Hennig-Schmidt (2011). "The Framing of Games and the Psychology of Play." Games and Economic Behavior **73**: 459-478.
- Engel, Christoph (2015a). "Scientific Disintegrity as a Public Bad." Perspectives on Psychological Science **10**(3): 361-379.

Engel, Christoph (2015b). "Tacit Collusion. The Neglected Experimental Evidence." Journal of Empirical Legal Studies **12**: 537-577.

Engel, Christoph and Bettina Rockenbach (2024). "What Makes Cooperation Precarious?" Journal of Economic Psychology **101**: 102712.

Epley, Nicholas and Thomas Gilovich (2016). "The Mechanics of Motivated Reasoning." Journal of Economic Perspectives **30**(3): 133-140.

Etner, Johanna, Meglena Jeleva and Jean-Marc Tallon (2012). "Decision Theory under Ambiguity." Journal of Economic Surveys **26**(2): 234-270.

Fairfield, Joshua AT and Christoph Engel (2015). "Privacy as a Public Good." Duke Law Journal **65**: 385-457.

Falk, Armin, Ernst Fehr and Urs Fischbacher (2008). "Testing Theories of Fairness - Intentions Matter." Games and Economic Behavior **62**: 287-303.

Fehr, Ernst and Klaus M. Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation." Quarterly Journal of Economics **114**: 817-868.

Fehr, Ernst and Klaus M. Schmidt (2006). "The Economics of Fairness, Reciprocity and Altruism. Experimental Evidence and New Theories." Handbook on the Economics of Giving, Reciprocity and Altruism **1**: 615-691.

Fershtman, Chaim and Shmuel Nitzan (1991). "Dynamic Voluntary Provision of Public Goods." European Economic Review **35**: 1057-1067.

Fischbacher, Urs (2007). "Z-Tree. Zurich Toolbox for Ready-Made Economic Experiments." Experimental Economics **10**: 171-178.

Fischbacher, Urs and Simon Gächter (2010). "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Good Experiments." American Economic Review **100**: 541-556.

Fischbacher, Urs, Simon Gächter and Ernst Fehr (2001). "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." Economics Letters **71**: 397-404.

Frederick, Shane, George Loewenstein and Ted O'Donoghue (2002). "Time Discounting and Time Preference. A Critical Review." Journal of Economic Literature **40**(2): 351-401.

Fudenberg, Drew, David K. Levine and Eric Maskin (1994). "The Folk Theorem with Imperfect Public Information." Econometrica **62**(5): 997-1039.

Fudenberg, Drew and Eric Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." Econometrica **54**: 533-554.

Gächter, Simon, Elke Renner and Martin Sefton (2008). "The Long-Run Benefits of Punishment." Science **322**: 1510-1510.

Ghidoni, Riccardo, Giacomo Calzolari and Marco Casari (2017). "Climate Change: Behavioral Responses from Extreme Events and Delayed Damages." Energy Economics **68**: 103-115.

Guilfoos, Todd, Haoran Miao, Simona Trandafir and Emi Uchida (2019). "Social Learning and Communication with Threshold Uncertainty." Resource and Energy Economics **55**: 81-101.

Haisley, Emily C. and Roberto A. Weber (2010). "Self-Serving Interpretations of Ambiguity in Other-Regarding Behavior." Games and Economic Behavior **68**(2): 614-625.

Hardin, Garrett (1968). "The Tragedy of the Commons." Science **162**: 1243-1248.

Isaac, R. Mark and James M. Walker (1988). "Group Size Effects in Public Goods Provision. The Voluntary Contributions Mechanism." Quarterly Journal of Economics **103**: 179-199.

Isaac, R. Mark, James M. Walker and Arlington W. Williams (1994). "Group Size and the Voluntary Provision of Public Goods. Experimental Evidence Utilizing Large Groups." Journal of Public Economics **54**: 1-36.

Isaac, R. Mark, James Walker and Susan H. Thomas (1984). "Divergent Evidence of Free Riding. An Experimental Examination of Possible Explanations." Public Choice **43**: 113-149.

Keizer, Kees, Siegwart Lindenberg and Linda Steg (2008). The Spreading of Disorder. Science **322**: 1681-1685.

Khadjavi, Menusch and Andreas Lange (2015). "Doing Good or Doing Harm: Experimental Evidence on Giving and Taking in Public Good Games." Experimental Economics **18**: 432-441.

Konow, James (2000). "Fair Shares. Accountability and Cognitive Dissonance in Allocation Decisions." American Economic Review **90**(4): 1072-1091.

Konow, James (2003). "Which Is the Fairest One of All? A Positive Analysis of Justice Theories." Journal of Economic Literature **41**: 1186-1237.

Kossioris, G., M. Plexousakis, A. Xepapadeas, A. de Zeeuw and K.-G. Mäler (2008). "Feedback Nash Equilibria for Non-Linear Differential Games in Pollution Control." Journal of Economic Dynamics and Control **32**: 1312-1331.

Kreps, David M., Paul R. Milgrom, John Roberts and Robert B. Wilson (1982). "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." Journal of Economic Theory **27**: 245-252.

Krupka, Erin and Roberto A. Weber (2009). "The Focusing and Informational Effects of Norms on Pro-Social Behavior." Journal of Economic Psychology **30**: 307-320.

Kunda, Ziva (1990). "The Case for Motivated Reasoning." Psychological Bulletin **108**: 480-498.

Laibson, David (1997). "Golden Eggs and Hyperbolic Discounting." Quarterly Journal of Economics **112**: 443-477.

Laury, Susan K., James Walker and Arlington W. Williams (1999). "The Voluntary Provision of a Pure Public Good with Diminishing Marginal Returns." Public Choice **99**: 139-160.

Ledyard, John O. (1995). Public Goods. A Survey of Experimental Research. The Handbook of Experimental Economics. J. H. Kagel and A. E. Roth. Princeton, NJ, Princeton University Press: 111-194.

Moghim, Sanaz and Roja Kaveh Garna (2019). "Countries' Classification by Environmental Resilience." Journal of Environmental Management **230**: 345-354.

Moxnes, Erling and Eline Van der Heijden (2003). "The Effect of Leadership in a Public Bad Experiment." Journal of Conflict Resolution **47**: 773-795.

Murphy, Ryan O, Kurt A Ackermann and Michel Handgraaf (2011). "Measuring Social Value Orientation." Judgment and Decision Making **6**(8): 771-781.

Nelson, Donald R, W Neil Adger and Katrina Brown (2007). "Adaptation to Environmental Change: Contributions of a Resilience Framework." Annual Review of Environmental Resources **32**(1): 395-419.

Noussair, Charles and Cindy Soo (2008). "Voluntary Contributions to a Dynamic Public Good: Experimental Evidence." Economics Letters **98**(1): 71-77.

O'Donoghue, Ted and Matthew Rabin (2015). "Present Bias. Lessons Learned and to Be Learned." American Economic Review **105**(5): 273-279.

O'Donoghue, Ted and Matthew Rabin (1999). "Doing It Now or Later." American Economic Review **89**(1): 103-124.

Pevnitskaya, Svetlana and Dmitry Ryvkin (2013). "Environmental Context and Termination Uncertainty in Games with a Dynamic Public Bad." Environment and Development Economics **18**(1): 27-49.

Pevnitskaya, Svetlana and Dmitry Ryvkin (2022). "Strategy Adjustments in Games with a Dynamic Public Bad: Experimental Evidence." Review of Behavioral Economics **9**(2): 173-184.

Rabin, Matthew (1993). "Incorporating Fairness into Game Theory and Economics." American Economic Review **83**: 1281-1302.

Rammstedt, Beatrice and Oliver P. John (2007). "Measuring Personality in One Minute or Less. A 10-Item Short Version of the Big Five Inventory in English and German." Journal of Research in Personality **41**: 203-212.

Sarin, Rakesh K and Martin Weber (1993). "Effects of Ambiguity in Market Experiments." Management Science **39**(5): 602-615.

Sonnemans, Joep, Arthur Schram and Theo Offerman (1998). "Public Good Provision and Public Bad Provision. The Effect of Framing." Journal of Economic Behavior & Organization **34**: 143-161.

Susser, Daniel, Beate Roessler and Helen Nissenbaum (2019). "Online Manipulation: Hidden Influences in a Digital World." Georgia Law Technology Review **4**: 1-46.

Ulrich, Martin (2021). Politische Ideologien (Polid).

Wirl, Franz (1996). "Dynamic Voluntary Provision of Public Goods. Extension to Nonlinear Strategies." European Journal of Political Economy **12**: 555-560.

Zelmer, Jennifer (2003). "Linear Public Goods. A Meta-Analysis." Experimental Economics **6**: 299-310.

**Out of Sight is Out of Mind?**  
**Experimentally Testing a Public Bad That Only Materializes Gradually**  
**Appendix A**

**Table A1 – Summary Statistics: Balancing Table between Decay Rate Treatments**

Decay Rate	Low Decay (0.1)			High Decay (0.9)			Test
Variable	N	Mean	SD	N	Mean	SD	
age	126	26.3	8.82	126	26.2	9.49	F=0.023
gender_cat	126			126			X2=2.907
... Male	55	43.7%		50	39.7%		
... Diverse/Not Disclosed	9	7.1%		4	3.2%		
... Female	62	49.2%		72	57.1%		
work_cat	126			126			X2=0.043
... No	12	9.5%		14	11.1%		
... Yes	114	90.5%		112	88.9%		
experience_experimental	126			126			X2=0.455
... No Experience	20	15.9%		18	14.3%		
... 1-2 Times	24	19%		21	16.7%		
... 3-5 Times	28	22.2%		30	23.8%		
... More than 5	54	42.9%		57	45.2%		
siblings	126	1.79	1.43	126	1.6	1.38	F=1.158
semester	104	6.34	4.74	108	6.07	4.49	F=0.171
svo_type	126			126			X2=2.017
... Altruism	0	0%		1	0.8%		
... Prosocial	92	73%		84	66.7%		
... Individualist	34	27%		41	32.5%		
PCA_krupka	126	-0.245	2.59	126	0.245	2.57	F=2.274
TimeOKControl_Total	126	865	347	126	888	296	F=0.321
QuestFalse_Total	126	9.37	8.78	126	10.3	8.05	F=0.702
BFI_neurotic	126	3	1.12	126	2.9	0.978	F=0.56
BFI_opennes	126	2.27	1	126	2.24	1.01	F=0.062
BFI_extravers	126	3.05	1	126	2.66	0.981	F=9.662***
BFI_agreeable	126	2.83	0.749	126	2.86	0.88	F=0.12
BFI_conscient	126	2.59	0.866	126	2.38	0.823	F=3.905**
POLID_Conservatism	126	-0.602	0.462	126	-0.504	0.455	F=2.866*
POLID_EconLibertarianism	126	-0.545	0.612	126	-0.475	0.631	F=0.798
POLID_Socialism	126	0.386	0.597	126	0.309	0.641	F=0.971
POLID_Liberalism	126	0.761	0.503	126	0.67	0.51	F=2.026
Concerns_Privacy	126	-0.136	0.401	126	-0.0948	0.397	F=0.674
solo_treat	126			126			X2=0
... 0	96	76.2%		96	76.2%		
... 1	30	23.8%		30	23.8%		

Statistical significance markers: \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

**Table A2 – Summary Statistics: Balancing Table between Single and Group Treatments**

Type of Interaction	Group Treatment (Human)			Solo Treatment (Computer)			Test
Variable	N	Mean	SD	N	Mean	SD	
age	192	26.4	8.84	60	25.9	10.1	F=0.159
gender_cat	192			60			X2=1.13
... Male	77	40.1%		28	46.7%		
... Diverse/Not Disclosed	11	5.7%		2	3.3%		
... Female	104	54.2%		30	50%		
work_cat	192			60			X2=0
... No	20	10.4%		6	10%		
... Yes	172	89.6%		54	90%		
experience_experimental	192			60			X2=5.958
... No Experience	29	15.1%		9	15%		
... 1-2 Times	30	15.6%		15	25%		
... 3-5 Times	41	21.4%		17	28.3%		
... More than 5	92	47.9%		19	31.7%		
siblings	192	1.61	1.4	60	1.93	1.4	F=2.364
semester	159	6.28	4.74	53	5.96	4.2	F=0.192
svo_type	192			60			X2=0.704
... Altruism	1	0.5%		0	0%		
... Prosocial	132	68.8%		44	73.3%		
... Individualist	59	30.7%		16	26.7%		
PCA_krupka	192	0.184	2.62	60	-0.588	2.39	F=4.12**
BFI_neurotic	192	2.93	1.08	60	3.01	0.946	F=0.238
BFI_opennes	192	2.26	1.03	60	2.23	0.932	F=0.033
BFI_extravers	192	2.79	1.01	60	3.05	0.982	F=3.014*
BFI_agreeable	192	2.86	0.838	60	2.78	0.744	F=0.424
BFI_conscient	192	2.48	0.907	60	2.52	0.638	F=0.101
POLID_Conservativm	192	-0.556	0.476	60	-0.544	0.409	F=0.034
POLID_EconLibertarianism	192	-0.546	0.635	60	-0.394	0.566	F=2.762*
POLID_Socialism	192	0.377	0.607	60	0.252	0.652	F=1.876
POLID_Liberalism	192	0.725	0.523	60	0.685	0.457	F=0.277
Concerns_Privacy	192	-0.109	0.413	60	-0.135	0.351	F=0.188
decay_high	192			60			X2=0
... 0	96	50%		30	50%		
... 1	96	50%		30	50%		

Statistical significance markers: \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

*Table A2*  
*Summary Statistics: Balancing Table between Single and Group Treatments*

**Table A3 – Group Contribution with Period Squared**

<i>Predictors</i>	<b>Model 3 (Period Squared)</b>		
	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>
(Intercept)	5.779 ***	0.705	<b>&lt;0.001</b>
decay high [1]	0.938	0.997	0.347
period	0.284 ***	0.062	<b>&lt;0.001</b>
period squared	-0.002	0.003	0.511
decay high [1] × period	-0.033	0.087	0.704
decay high [1] × period squared	0.004	0.004	0.276
<b>Random Effects</b>			
$\sigma^2$	13.717		
$\tau_{00}$ subject_uid:group_uid	6.883		
$\tau_{00}$ group_uid	8.307		
ICC	0.525		
N subject_uid	192		
N group_uid	48		
Observations	3840		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.091 / 0.569		

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

Linear Mixed Effects Model

Standard errors for choices nested in individuals nested in groups

[illegible]





Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.107 / 0.573	0.098 / 0.561	0.129 / 0.569	0.093 / 0.568	0.093 / 0.571	0.096 / 0.568	0.090 / 0.569	0.160 / 0.566
---	---------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------

*\* p<0.05   \*\* p<0.01   \*\*\* p<0.001*

Linear Mixed Effects Model

Standard errors for choices nested in individuals nested in groups

[illegible]

POLID EconLibertarianism	0.408	0.439						-0.037	0.453
POLID Liberalism	0.832	0.640						0.348	0.615
BFI agreeable				0.275	0.263			0.095	0.272
BFI extravers				-0.235	0.241			-0.042	0.236
BFI conscient				-0.072	0.259			-0.210	0.269
BFI neurotic				-0.071	0.227			-0.011	0.250
Concerns Privacy						1.049	0.544	0.700	0.544
TimeOKControl Total								-0.000	0.001

# Random Effects

$\sigma^2$	13.714	13.714	13.714	13.714	13.714	13.714	13.714	13.714
$\tau_{00}$	6.804	6.961	5.597	6.991	6.978	6.755	6.905	5.829
	subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid
	8.197	7.516	8.429	8.099	8.317	8.291	8.387	7.010
	group_uid	group_uid	group_uid	group_uid	group_uid	group_uid	group_uid	group_uid
ICC	0.522	0.514	0.506	0.524	0.527	0.523	0.527	0.484
N	192	192	192	192	192	192	192	192
	subject_uid	subject_uid	subject_uid	subject_uid	subject_uid	subject_uid	subject_uid	subject_uid
	48	48	48	48	48	48	48	48
	group_uid	group_uid	group_uid	group_uid	group_uid	group_uid	group_uid	group_uid
Observations	3840	3840	3840	3840	3840	3840	3840	3840
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.107 / 0.574	0.098 / 0.561	0.130 / 0.570	0.094 / 0.569	0.094 / 0.572	0.097 / 0.569	0.091 / 0.570	0.161 / 0.567

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

Linear Mixed Effects Model

Standard errors for choices nested in individuals nested in groups

[illegible]

PCA krupka	0.162	0.089							0.249 **	0.090
svo individualist [Yes]			2.423 ***	0.433					2.480 ***	0.476
POLID Conservativm				0.539	0.599				0.363	0.571
POLID EconLibertarianism				0.408	0.439				-0.037	0.453
POLID Liberalism				0.832	0.640				0.348	0.615
BFI agreeable						0.275	0.263		0.095	0.272
BFI extravers						-0.235	0.241		-0.042	0.236
BFI conscient						-0.072	0.259		-0.210	0.269
BFI neurotic						-0.071	0.227		-0.011	0.250
Concerns Privacy								1.049	0.544	
TimeOKControl Total									-0.000	0.001
									0.000	0.001

### Random Effects

$\sigma^2$	13.717	13.717	13.717	13.717	13.717	13.717	13.717	13.717								
$\tau_{00}$	6.804	6.961	5.597	subject_uid:group_uid	6.991	6.977	6.754	6.904	5.829							
	subject_uid:group_uid	subject_uid:group_uid			subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid	subject_uid:group_uid							
	8.196	group_uid	7.516	group_uid	8.429	group_uid	8.099	group_uid	8.317	group_uid	8.291	group_uid	8.387	group_uid	7.010	group_uid
ICC	0.522	0.513	0.506		0.524	0.527	0.523	0.527	0.483							
N	192	subject_uid	192	subject_uid	192	subject_uid	192	subject_uid	192	subject_uid	192	subject_uid	192	subject_uid	192	subject_uid
	48	group_uid	48	group_uid	48	group_uid	48	group_uid	48	group_uid	48	group_uid	48	group_uid	48	group_uid
Observations	3840	3840	3840		3840	3840	3840	3840	3840							
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.108 / 0.574	0.099 / 0.561	0.130 / 0.570		0.094 / 0.569	0.094 / 0.572	0.097 / 0.569	0.091 / 0.570	0.161 / 0.567							

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

Linear Mixed Effects Model

Standard errors for choices nested in individuals nested in groups

**Table A7 – Solo vs. Group Treatments with Decay and Period Effects**

<i>Predictors</i>	<b>Model 1</b>			<b>Model 2</b>			<b>Model 3</b>		
	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>
(Intercept)	8.605 ***	0.541	<0.001	8.490 ***	0.615	<0.001	5.923 ***	0.634	<0.001
solo treat [1]	-4.471 ***	0.655	<0.001	-4.207 ***	0.930	<0.001	-2.931 **	0.982	0.003
decay high [1]	0.990	0.650	0.128	1.221	0.869	0.160	0.600	0.896	0.503
solo treat [1] × decay high [1]				-0.529	1.315	0.687	-1.075	1.389	0.439
period							0.244 ***	0.015	<0.001
solo treat [1] × period							-0.122 ***	0.030	<0.001
decay high [1] × period							0.059 **	0.021	0.005
(solo treat [1] × decay high [1]) × period							0.052	0.043	0.224
<b>Random Effects</b>									
$\sigma^2$	16.194			16.194			13.888		
$\tau_{00}$	6.574	subject_uid:group_uid		6.579	subject_uid:group_uid		6.694	subject_uid:group_uid	
	7.138	group_uid		7.219	group_uid		7.218	group_uid	
ICC	0.459			0.460			0.500		
N	252	subject_uid		252	subject_uid		252	subject_uid	
	108	group_uid		108	group_uid		108	group_uid	
Observations	5040			5040			5040		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.115 / 0.521			0.116 / 0.523			0.181 / 0.591		

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

Linear Mixed Effects Model

Standard errors for choices nested in individuals nested in groups

**Table A8 – Selfish vs. Not Selfish Participants with Decay and Period Effects**

<i>Predictors</i>	<b>Model 1</b>			<b>Model 2</b>			<b>Model 3</b>		
	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>
(Intercept)	7.809 ***	0.657	<0.001	7.807 ***	0.669	<0.001	5.136 ***	0.693	<0.001
selfish	2.423 ***	0.433	<0.001	2.427 ***	0.622	<0.001	2.798 ***	0.710	<0.001
decay high [1]	1.095	0.913	0.231	1.097	0.952	0.249	0.379	0.987	0.701
selfish × decay high [1]				-0.008	0.868	0.993	0.226	0.989	0.819
period							0.254 ***	0.017	<0.001
selfish × period							-0.035	0.033	0.278
decay high [1] × period							0.068 **	0.025	0.006
(selfish × decay high [1]) × period							-0.022	0.045	0.621
<b>Random Effects</b>									
$\sigma^2$	16.365			16.365			13.704		
$\tau_{00}$	5.465	subject_uid:group_uid		5.507	subject_uid:group_uid		5.640	subject_uid:group_uid	
	8.429	group_uid		8.429	group_uid		8.429	group_uid	
ICC	0.459			0.460			0.507		
N	192	subject_uid		192	subject_uid		192	subject_uid	
	48	group_uid		48	group_uid		48	group_uid	
Observations	3840			3840			3840		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.051 / 0.487			0.051 / 0.487			0.130 / 0.571		

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

Linear Mixed Effects Model

Standard errors for choices nested in individuals nested in groups



**Table A9 – Time to respond to control questions as a predictor in Single Treat**

<i>Predictors</i>	<b>Model 1 Normative Choice Dummy</b>			<b>Model 2 Normative Choice Interaction</b>			<b>Model 3 Contribution period 20 omitted</b>			<b>Model 4 Contribution with Interaction period 20 omitted</b>		
	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>
(Intercept)	1.063 ***	0.126	<0.001	1.169 ***	0.152	<0.001	-0.290	1.074	0.787	-0.991	1.433	0.489
Minutes Control	-0.034 ***	0.007	<0.001	-0.040 ***	0.009	<0.001	0.281 ***	0.062	<0.001	0.297 ***	0.080	<0.001
decay high [1]	-0.175	0.089	0.050	-0.439	0.235	0.061				1.411	2.204	0.522
Minutes Control × decay high [1]				0.017	0.014	0.224				-0.034	0.128	0.790
<b>Random Effects</b>												
$\sigma^2$	0.079			0.079			10.787			10.787		
$\tau_{00}$	0.115 subject_uid			0.114 subject_uid			9.678 subject_uid			9.829 subject_uid		
ICC	0.591			0.589			0.473			0.477		
N	60 subject_uid			60 subject_uid			60 subject_uid			60 subject_uid		
Observations	1200			1200			1140			1140		
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.228 / 0.684			0.237 / 0.687			0.147 / 0.550			0.153 / 0.557		

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

Table A10 – Control Question Time on Contribution in Group

<i>Predictors</i>	<b>Model 1</b>			<b>Model 2</b>			<b>Model 3</b>			<b>Model 4</b>		
	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>
(Intercept)	6.221 ***	0.925	<b>&lt;0.001</b>	5.136 ***	0.905	<b>&lt;0.001</b>	4.630 ***	1.003	<b>&lt;0.001</b>	5.281 ***	1.104	<b>&lt;0.001</b>
decay high [1]	0.615	0.952	0.518	0.468	0.939	0.618	0.464	0.941	0.622	-1.357	1.678	0.419
period	0.244 ***	0.015	<b>&lt;0.001</b>	0.244 ***	0.015	<b>&lt;0.001</b>	0.244 ***	0.015	<b>&lt;0.001</b>	0.254 ***	0.043	<b>&lt;0.001</b>
MinutesControl	-0.022	0.046	0.639	0.007	0.042	0.862	0.042	0.052	0.416	-0.003	0.063	0.959
decay high [1] × period	0.059 **	0.021	<b>0.004</b>	0.059 **	0.021	<b>0.004</b>	0.059 **	0.021	<b>0.004</b>	0.171 **	0.065	<b>0.009</b>
svo individualist [Yes]				2.433 ***	0.438	<b>&lt;0.001</b>	3.963 **	1.373	<b>0.004</b>	2.440 ***	0.439	<b>&lt;0.001</b>
MinutesControl × svo individualist [Yes]							-0.110	0.093	0.240			
decay high [1] × MinutesControl										0.126	0.098	0.196
MinutesControl × period										-0.001	0.003	0.813
(decay high [1] × MinutesControl) × period										-0.008	0.004	0.078

**Random Effects**

$\sigma^2$	13.714	13.714	13.714	13.696
$\tau_{00}$	6.905 <small>subject_uid:group_uid</small>	5.643 <small>subject_uid:group_uid</small>	5.615 <small>subject_uid:group_uid</small>	5.673 <small>subject_uid:group_uid</small>
	8.387 <small>group_uid</small>	8.404 <small>group_uid</small>	8.465 <small>group_uid</small>	8.411 <small>group_uid</small>
ICC	0.527	0.506	0.507	0.507
N	192 <small>subject_uid</small>	192 <small>subject_uid</small>	192 <small>subject_uid</small>	192 <small>subject_uid</small>
	48 <small>group_uid</small>	48 <small>group_uid</small>	48 <small>group_uid</small>	48 <small>group_uid</small>
Observations	3840	3840	3840	3840
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.091 / 0.570	0.130 / 0.570	0.132 / 0.572	0.131 / 0.572

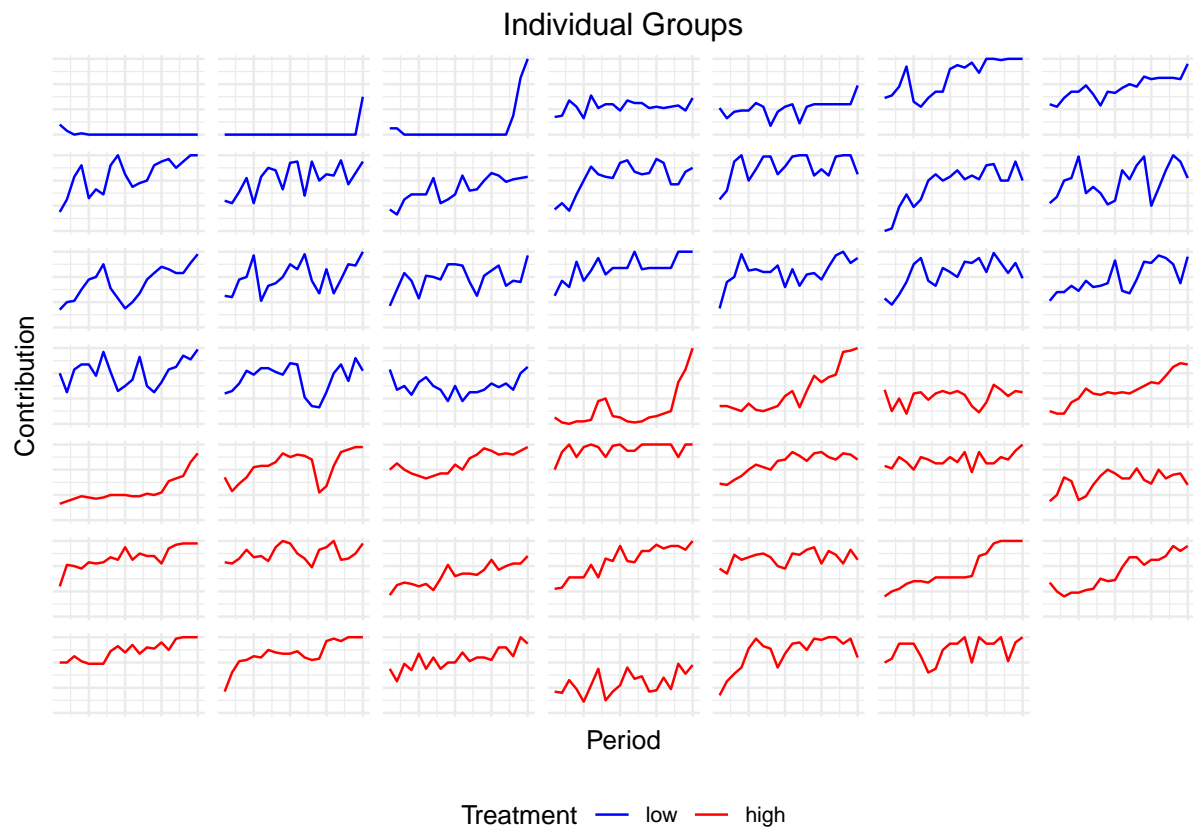
\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

**Table A11 – Learning from experiences in the Single Treatment**

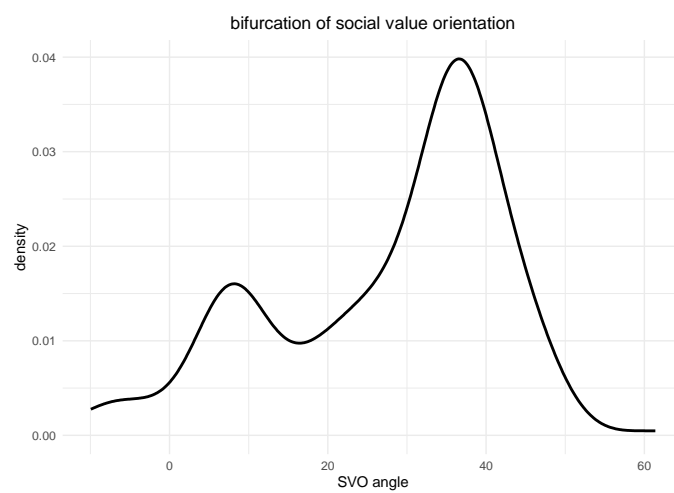
<i>Predictors</i>	<b>contribution</b>					
	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>	<i>Estimates</i>	<i>std. Error</i>	<i>p</i>
(Intercept)	4.662 ***	0.649	< <b>0.001</b>	4.810 ***	0.659	< <b>0.001</b>
decay high [1]	0.671	0.903	0.458	0.379	0.931	0.684
distneg1	-0.798	0.534	0.135	-1.410	0.756	0.062
distneg2	-0.932	0.538	0.084	-0.843	0.756	0.265
distneg3	-1.084 *	0.538	<b>0.044</b>	-1.343	0.756	0.076
distneg4	-1.033	0.538	0.055	-0.743	0.756	0.326
distneg5	-1.169 *	0.538	<b>0.030</b>	-1.710 *	0.756	<b>0.024</b>
distneg6	-0.016	0.538	0.976	-0.210	0.756	0.781
distneg7	-0.253	0.543	0.641	-0.977	0.756	0.197
distneg8	-0.874	0.543	0.108	-1.343	0.756	0.076
distneg9	-0.426	0.543	0.433	-1.277	0.756	0.092
distneg10	-0.992	0.557	0.075	-0.677	0.756	0.371
decay high [1] × distneg1				1.221	1.068	0.253
decay high [1] × distneg2				-0.185	1.078	0.864
decay high [1] × distneg3				0.522	1.078	0.628
decay high [1] × distneg4				-0.595	1.078	0.581
decay high [1] × distneg5				1.096	1.078	0.310
decay high [1] × distneg6				0.389	1.078	0.718
decay high [1] × distneg7				1.491	1.087	0.171
decay high [1] × distneg8				0.964	1.087	0.375
decay high [1] × distneg9				1.755	1.087	0.107
decay high [1] × distneg10				-0.723	1.119	0.518
<b>Random Effects</b>						
$\sigma^2$	15.565			15.593		

$\tau_{00}$	11.461 <sub>subject_uid</sub>	11.464 <sub>subject_uid</sub>
ICC	0.424	0.424
N	60 <sub>subject_uid</sub>	60 <sub>subject_uid</sub>
Observations	1200	1200
Marginal R <sup>2</sup> / Conditional R <sup>2</sup>	0.012 / 0.431	0.016 / 0.433
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$		

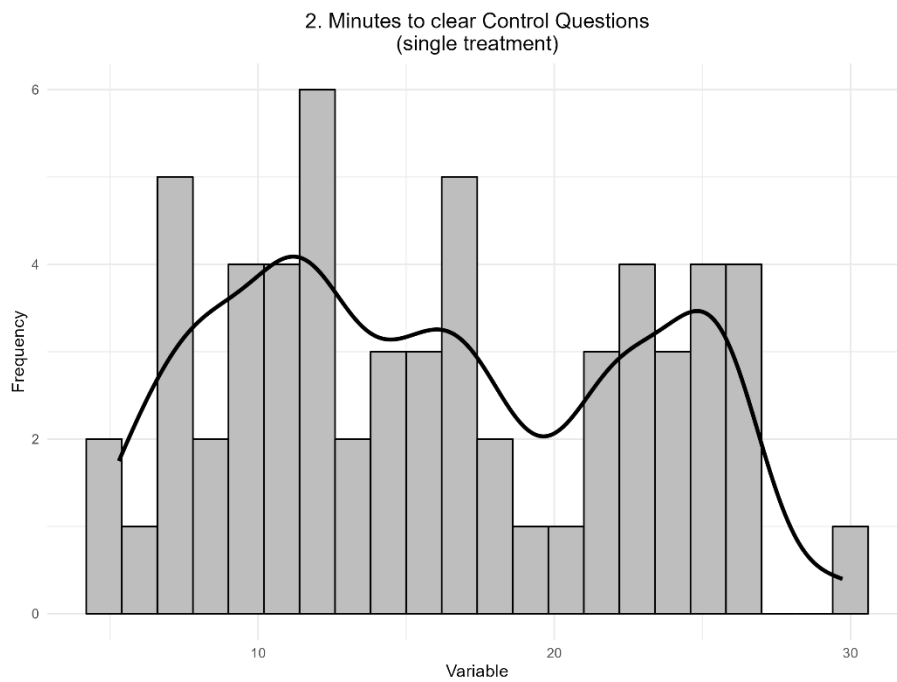
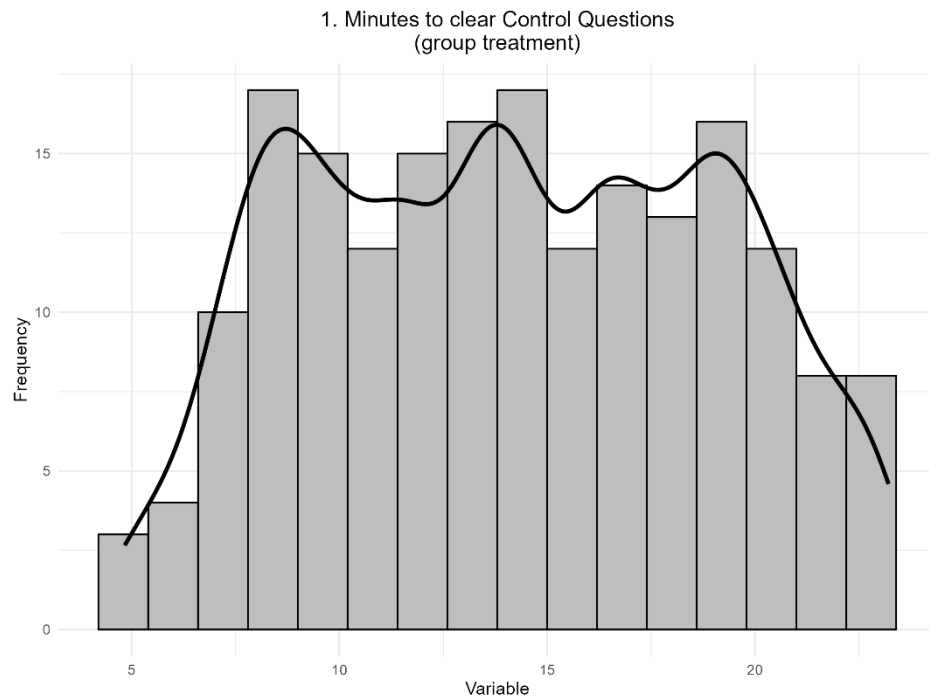
**Figure A1 – Group Treatments: Average Contributions per Group**



**Figure A2 – Bifurcation of Social Value Orientation**



**Figures A3.1 and A3.2 – Minutes need to respond to Control Questions**



**Out of Sight is Out of Mind?**  
**Experimentally Testing a Public Bad That Only Materializes Gradually**  
**Appendix B**

**I. English Translation of Experimental Instructions**

Below, you will find an English translation of the experimental instructions that were handed out to the participants in the beginning of the experiment. Parts of the instructions that differ between the High Decay Rate Treatment and the Low Decay Rate Treatment are indicated in **yellow highlight** and separated with a “|”. Text which was only presented in the Single Treatment is indicated in **green highlight**.

**Instructions**

Welcome to our experiment. During the experiment, communication with other participants is not allowed. Please turn off your mobile devices and put them out of reach. If you have any questions, please make yourself known by a clear hand signal. The test supervisor will try to answer your questions.

**Procedure**

The experiment consists of **three parts** and a questionnaire. The instructions for the first part can be found in this document. You will receive the instructions for the other two parts on the computer screen as soon as the first part is finished. Please read the instructions carefully and completely. Make your decision deliberately and only after all your questions have been answered. The three parts are independent of each other and your decisions in the previous parts have no influence on the later ones. We will explain separately for each part of the experiment how your payout will be calculated. At the end of the experiment, we will **pay out your total earnings in cash**. Your total earnings will result from the decisions you make during the experiment. We will treat all your choices confidentially.

**Part 1**

In this part of the experiment, you will be presented with the choice of investing in a project that has a detrimental effect on the group to which you are assigned. If you decide to invest in the project, you will receive a personal financial benefit. However, this decision also leads to disadvantages for you and your group members. You therefore face a social dilemma: for you personally, it is profitable to invest into the project. However, if all members of your group invest the same amount, this results in a disadvantage that is greater than the individual returns.



## General Information

In this part of the experiment, you form a group with three other group members. You can interact with them for 20 rounds. At the start of the experiment, you and each group member **receive a financial budget of 28.00 EUR**. This “budget” can increase or decrease over the course of the experiment. This will depend on the decisions made within your group.

You have the possibility to make investments within the Group. Each investment has advantages and disadvantages. The resulting advantages only benefit you personally. However, the disadvantages have long-term consequences for the entire group.

In this part of the experiment, we will be speaking about points. These points represent your ability to invest in a project. You can invest between 0 and 15 points in each round. Your group members can also invest up to 15 points in each round. Should you decide to invest points in this experiment, this will have the following consequences:

For every point you invest, 1 cent is deducted from your initial balance (**cost**). You will also receive 10 cents back for every point you invest (**benefit**).

### Example 1:

You invest 9 points. 9 cents are deducted; you receive  $10 \times 9 = 90$  cents back. The net amount you receive is therefore:  $- 9 \text{ cents} + (10 \times 9 \text{ cents}) = 81 \text{ cents}$ .

In addition, a **deduction** may be made in each round. The amount of the deduction depends on how many points you and your party members have invested in the current and previous rounds (**detriment**).

## Calculation of the Detriment

Each point that you or other group members invest, creates a detriment of the same amount. The effects of the detriment affect you and your group members equally.

### Example 2:

Round 1: You invest 9 points. Your group members invest 7, 14 and 10 points. In total,  $9+7+14+10=40$  points are invested in this round. The detriment in this round is 40.

The investments made in the current round and the corresponding detriment do not only affect the current round, but also future rounds. The detriment is reduced over the course of several rounds. The current detriment is gradually reduced from one round to the next by a factor of **0.9 | 0.1**.

In other words, the detriment created in the current round is multiplied by **0.9 | 0.1** and carried over to the following round. Further detriment is added in the following round if more points are invested then. For the next subsequent round, the new disadvantage is then again multiplied by **0.9 | 0.1** and carried over. See Figure 1 for an explanation of this mechanism.

## Calculation of the Deduction

The deduction in a round is calculated by **multiplying the current detriment by 0.3 | 2.7**. The result reduces the profit per round for all group members. The following example illustrates the calculation of detriment and deduction:

	Rounds								
	1	2	3	4	5 to 20 ...				
Round 1: 40 Points	40	$\times 0.9$	36	$\times 0.9$	32.4	$\times 0.9$	29.16	$\times 0.9$	...
Round 2: 24 Points		+	24	$\times 0.9$	21.6	$\times 0.9$	19.44	$\times 0.9$	...
Round 3: 36 Points			+	36	$\times 0.9$	32.4	$\times 0.9$	...	
Round 4: 39 Points				+	39	$\times 0.9$	...		
...									
Sum of Detriments:	= 40	= 60	= 90	= 120					
$\times 0.3 =$	$\times 0.3$	$\times 0.3$	$\times 0.3$	$\times 0.3$					
Deduction per Round:	= 12	= 18	= 27	= 36					

Figure 1 (High Decay) ↑ | ↓ Figure 1 (Low Decay)

	Rounds								
	1	2	3	4	5 to 20 ...				
Round 1: 40 Points	40	$\times 0.1 \rightarrow$	4	$\times 0.1 \rightarrow$	0.4	$\times 0.1 \rightarrow$	0.04	$\times 0.1 \rightarrow$	...
Round 2: 24 Points		+	24	$\times 0.1 \rightarrow$	2.4	$\times 0.1 \rightarrow$	0.24	$\times 0.1 \rightarrow$	...
Round 3: 36 Points			+	36	$\times 0.1 \rightarrow$	3.6	$\times 0.1 \rightarrow$	...	
Round 4: 39 Points				+	39	$\times 0.1 \rightarrow$	...		
...									
Sum of Detriments:	= 40	= 28	= 38.8	= 42.88					
$\times 2.7 =$	$\times 2.7$	$\times 2.7$	$\times 2.7$	$\times 2.7$					
Deduction per Round:	= 108	= 75.6	= 104.76	= 115.78					

## Calculation of Profit per Round

The profit per round is composed of the elements explained above: The cost of the investment reduces the profit each round, the benefits increase it, the deduction reduces it. This results in the following formula:

		Cost:		Benefit:		Deduction:
Profit per round	=	—	1 Cent per invested Point in the current round (0 to 15 Cent)	+	10 Cent per invested Point in the current round (0 to 150 Cent)	— $0.3 \mid 2.7 \times$ current detriment

### Long-term effects of Investment Decisions

You and your group members can make investment decisions for 20 rounds. The long-term effects of an investment on your payoff are the same, regardless of the round in which you or other group members invest points. Your payoff is calculated on the assumption that the experiment would have continued indefinitely. As a result, for every point invested within your group, **a total of 3 cents will be deducted** from you and each of your group members.

This long-term consequence is taken into account for every point invested within the 20 rounds. All deductions that would only be incurred after round 20 will be **deducted in a one-off payment after completing round 20**.

#### Example 3:

Round 1: One point is invested.

A deduction of 2.64 cents is made until the end of round 20. To ensure that a total deduction of 3 cents is made for the investment of the point, 0.36 cents will be deducted at the end of Part 1 in the one-time payment. | A deduction of 3 cents is made until the end of round 20. The total deduction of 3 cents is realized during the experiment, so that no further amount is deducted at the end of Part 1.

Round 20: One point is invested. A deduction of  $0.3 \mid 2.7$  cents is made until the end of round 20. To ensure that a total deduction of 3 cents is made for the investment of the point,  $2.7 \mid 0.3$  cents will be deducted at the end of Part 1 in the one-time payment.

### Behavior of the Computerized Group Members

In this part of the experiment, you do not interact with other participants. Instead, the behavior of your group members is controlled by a computer. The computer's choices solely depend on your behavior: The computer-controlled group members **always imitate your behavior from the previous round**.

#### Example 4:

Round 10: You invest 5 points.

Round 11: Each of your computerized group members also invests 5 points. Your group members therefore invest a total of 15 points.

As your computerized group members cannot yet imitate your behavior in the first round, their investments are determined randomly in round 1.

### Control Questions

In the experiment, we will automatically calculate the round profit for you. However, we will ask you to carry out some of these calculations yourself in control questions so that that you can familiarize yourself with the consequences of your decisions.

If you are asked to enter an amount with a decimal place, please use a period "." instead of a comma. If a negative value results, please place a minus (-) in front of the figure.

Please raise your hand if you have a question. The test supervisor will then try to answer your questions.

## II. English Translation of Control Questions

We have used an extensive set of control questions to ensure that participants have understood the core mechanics of our experiment. These control questions were programmed in an instructive manner: When a question is answered correctly a brief confirmation message is displayed and the next sub-question appears. When a question is answered incorrectly, our program provided a hint, pointing towards the relevant aspects of our mechanism without giving away the correct answer. Sub-questions are shown sequentially.

Below, we provide an English translation of the complete control questions for the High Decay Rate Treatment ( $\delta = 0.9$ ). For the sake of conciseness, we only illustrate one exemplary set of control questions. The same control questions were used in all treatments. There are only marginal differences in wording between treatments. It is either 'group members' (Group Treatment) or 'computerized group members' (Single Treatment). The multiplier changes in accordance with the decay rate. Please note that Control Question 5 was only included in the Single Treatment. Please see our preregistration materials online for screenshots of the whole experiment.

### Legend

- W: **hint after wrong answer**
- C: **confirmation message and explanation after correct answer**

### **Control Question 1:**

Assume that you are in the first round. You have invested 10 points. None of your group members have invested points in this round, so your total contribution is 10 points.

#### **1.1 How high are the costs resulting from your decision?**

W: **Unfortunately, your answer is wrong. Every point you invest costs 1 cent.**

C: **Each point invested incurs a cost of 1 cent. Costs have a negative effect on the round profit**

#### **1.2 What is your advantage?**

W: **Unfortunately, your answer is wrong. You will receive 10 cents for every point invested.**

C: **Each point invested gives you an advantage of 10 cents.**

#### **1.3 What is the current detriment?**

W: **Unfortunately, your answer is wrong. Each point invested in the group creates a detriment of the same amount in the same round.**

C: **The detriment here corresponds to the points you have invested. There is no detriment from previous rounds, as this is round 1.**

#### **1.4 How high is the deduction?**

W: **Unfortunately, your answer is wrong. The deduction corresponds to 0.3 times the detriment.**

C: **The deduction is calculated by multiplying the current detriment by 0.3. The deduction has a negative effect on the round profit.**

#### **1.5 Please calculate your personal winnings for this round from the above information:**

W: **Unfortunately, your answer is wrong. Take costs, benefits and deductions into account. Pay attention to the signs.**

C: **Your round profit results from the formula: -cost + advantage -deduction**

## Control Question 2:

Assume again that you are in the first round. You have invested 10 points. Your group members have invested 8, 10 and 12 points. A total of 40 points have therefore been invested.

2.1 How high are the costs resulting from your decision?

W: Unfortunately, your answer is wrong. Each point you invest costs 1 cent.

C: Each point invested incurs a cost of 1 cent. Costs have a negative effect on the round profit.

2.2 How high is your advantage?

W: Error message: Your answer is incorrect! You will receive 10 cents for every point invested.

C: Each point invested creates an advantage of 10 cents.

2.3 What is your detriment?

W: Unfortunately, your answer is wrong. Each point invested in the group creates a detriment of the same amount in the same round.

C: The detriment corresponds to the sum of all points invested in your group. There is no detriment from previous rounds, as this is round 1.

2.4 How high is the deduction?

W: Unfortunately, your answer is wrong. The deduction corresponds to 0.3 times the detriment.

C: The deduction is calculated by multiplying the current detriment by 0.3. The deduction has a negative effect on the round profit.

2.5 Please calculate your personal winnings for this round from the above information:

W: Unfortunately, your answer is wrong. Take costs, benefits and deductions into account. Pay attention to the signs.

C: Your round profit results from the formula:  $- \text{cost} + \text{advantage} - \text{deduction}$

## Control Question 3:

In the instructions, we explained how the investment of points affects the overall course of the experiment: The detriment resulting from invested points is reduced over time. It is multiplied by 0.9 for the detriment of the following round. To calculate the deduction in a round, the total detriment of a round is multiplied by 0.3. Assume a total of 40 points are invested in round 1. No further points are invested in rounds 2 and 3.

3.1 What is the detriment in round 1?

W: Unfortunately, your answer is wrong. You are only asking about detriment, not advantage or cost. Each point invested in the group creates a detriment of the same magnitude in the same round.

C: All points invested in round 1 are fully taken into account as detriment.

3.2 Assume that no further points were invested in round 2. How high is the detriment in round 2?

W: Unfortunately, your answer is wrong. The detriment from round 1 is reduced by a factor of 0.9.

C: The detriment from round 1 is multiplied by a factor of 0.9. Nothing was invested in round 2, so no further detriment is added.

3.3 Assume that no further points were invested in rounds 2 and 3. What is the detriment in round 3?

W: Unfortunately, your answer is wrong. The detriment from round 2 is reduced by a factor of 0.9.

C: The detriment from round 2 is again multiplied by a factor of 0.9. Once again, nothing was invested in round 3, so no further detriment is added.

### 3.4 Now calculate your personal profit for round 3 from the above results:

W: Unfortunately, your answer is incorrect! Your personal costs and benefits are based on the points you invested in the current round (0 points were invested in round 3). The deduction results from the detriment of the current round. Please note that a minus (-) must be placed in front of the number in the case of a negative profit

C: Since neither you nor your group members have invested 3 points in round 3, only the deduction resulting from the current detriment ( $0.3 \times 32.40$ ) is relevant for your round win.

### Control Question 4:

4.1 You are in round 1 and your group members have invested 12 points. Assume that no other points have been invested. How many deductions do you personally incur as a result of this behavior over the course of this experiment (20 rounds + one-off payment at the end)?

Hint: In this control question, you do not need to manually calculate the development of the detriments over 20 rounds. You only need to know the long-term effects of your investment decisions. If necessary, please refer to the bottom section on page 3 of the instructions.

W: Unfortunately, your answer is wrong. In the long term, every point invested generates a deduction of 3 cents.

C: Each point invested results in a long-term deduction of 3 cents per participant.

4.2 You are in round 17 and your group members have invested 12 points. Assume that no other points have been invested. How many deductions do you personally incur as a result of this behavior over the course of this experiment (20 rounds + one-off payment at the end)?

Hint: In this control question, you do not need to manually calculate the development of the detriments over 20 rounds. You only need to know the long-term effects of your investment decisions. If necessary, please refer to the bottom section on page 3 of the instructions.

W: Unfortunately, your answer is wrong. In the long term, every point invested generates a deduction of 3 cents.

C: Each point invested results in a long-term deduction of 3 cents per participant.

### Control Question 5:

If you decide to invest 11 points in round 6, how many points will be invested by your computerized group members in round 7? Investment of the group members (in total) in round 7:

W: Unfortunately, your answer is wrong. Every computerized member imitates your behavior from the previous round.

C: If you invest 11 points, each computerized group member will also invest 11 points in the following round.