Heterogeneous groups overcome the diffusion of responsibility problem in social norm enforcement

Wojtek Przepiorka[1], **Andreas Diekmann**[2,3]*

1 Department of Sociology/ICS, Utrecht University, Utrecht, Netherlands, **2** Department of Humanities, Social and Political Sciences, ETH Zurich, Zurich, Switzerland, **3** Institute of Advanced Studies, Berlin, Germany

* diekmann@soz.gess.ethz.ch

Abstract

Social norms promote cooperation in everyday life because many people are willing to negatively sanction norm breakers at a cost to themselves. However, a norm violation may persist if only one person is required to sanction the norm breaker and everyone expects someone else to do it. Here we employ the volunteer's dilemma game (VOD) to model this diffusion of responsibility in social norm enforcement. The symmetric VOD is a binary choice game in which all actors have the same costs of and benefits from cooperation and only one actor's cooperation is required to provide the collective good for the group. The asymmetric VOD differs from the symmetric VOD in one (strong) actor having lower costs of cooperation. In a laboratory experiment, we find that, in line with the diffusion of responsibility hypothesis, subjects' propensities to sanction the norm breaker decrease with group size in the symmetric VOD. In the asymmetric VOD, groups tacitly coordinate on the strong subject to sanction the norm breaker alone. Although at first lower in larger groups, strong subjects' sanctioning rates increase over time and reach equally high levels across different group sizes. Our results show that heterogeneous groups can be more effective in achieving norm compliance than groups of all equals because they naturally evade diffusion of responsibility in social norm enforcement.

Introduction

Cooperation problems studied by biologists, economists and other social scientists are often non-linear in nature [1,2,3,4,5]. In game theory, step-level collective goods games account for a specific type of non-linearity in the production function of collective goods [6,7,8,9]. The volunteers dilemma (VOD) is a prominent example. The VOD is a binary choice game in which a single actor's cooperation is necessary and sufficient to provide the collective good. Cooperators and free riders alike benefit from the collective good, but if everybody defects, no one receives anything [10]. The VOD and its extensions have been applied to a large number of collective good problems in such diverse fields as behavioural economics [11], biology [12],

computer science [13,14], sociology [15,16,17], and anthropology [18]. Here we show how the VOD can be applied to study diffusion of responsibility in social norm enforcement.

Social psychologists have extensively studied the causes of bystander intervention in emergency situations and diffusion of responsibility (aka bystander effect) has been of particular interest in this strand of research [19,20]. The diffusion of responsibility is a mechanism explaining the decreasing inclination of bystanders to help in an emergency situation the more other bystanders are present. In other words, diffusion of responsibility describes a group size effect by which the likelihood that the collective good is produced decreases with group size [10]. Research on social dilemmas [21,22,23] has produced a vast body of literature that investigates the effect of group size on cooperation and the production of collective goods [24,25,26,27,28]. Although the evidence is mixed, the literature is less equivocal in its concluding that the production function of a collective good is a major factor driving the group size effect.

Since Darley and Latané [19] corroborated the diffusion of responsibility effect empirically five decades ago, numerous studies have replicated it in various fields of application. Research has established several mechanisms involved in helping and other type of volunteering situations [29] and has explored various characteristics attenuating or strengthening diffusion of responsibility (for systematic reviews see [20,30].

Several authors proposed general models describing and explaining diffusion of responsibility. For example, Latané [31] (see also [32]) proposed an inverse power function to describe the group size effect while Piliavin et al. [33] suggested modelling the helping decision by a cost-benefit model. Although valuable steps forward, both these and other modelling efforts largely neglected the strategic nature of volunteering situations and game theoretic approaches to the modelling of the diffusion of responsibility effect. In what follows, we bring together psychology and game theory by modelling the volunteering situation as VOD and show how small variations in actors' game payoffs can lead to an *endogenous* solution of the diffusion of responsibility problem. We argue and show that this solution can be applied to the diffusion of responsibility problem in social norm enforcement.

The model

The basic VOD is a binary choice, symmetric *n*-person game in which a single actor's cooperation is sufficient to provide the collective good *U* at a cost *K*. Cooperators and free riders alike benefit from the collective good. However, if everybody free rides, no one receives anything [10] (Fig 1).

In the symmetric VOD, the costs and benefits from choosing C are the same for all actors (i.e. $U_i = U_j$ and $K_i = K_j \ \forall \ i \neq j$). The symmetric VOD has *n* asymmetric pure strategy Nash equilibria, in which one actor chooses C and all other actors choose D. A pure strategy equilibrium is, however, difficult to coordinate on tacitly in a one-shot game. The symmetric VOD also has a symmetric mixed strategy equilibrium, in which all actors have the same probability to choose C: $p_i^* = 1 - \sqrt[n-1]{K_i/U_i}$ [9]. The unique symmetric Nash equilibrium implies a monotonically decreasing function of the probability of cooperation (*p*) with group size (*n*). We call this property of the symmetric VOD the "diffusion of responsibility" or "bystander" effect [19] (see also [34]). A slight variation of the model allows for individual heterogeneity, i.e. cooperation costs (K_i) or benefits (U_i) vary across actors. The resulting asymmetric VOD has very different strategic properties.

In the asymmetric VOD with one "strong" actor *i* and *n*– 1 "weak" actors *j*, $U_i > U_j$ and/or $K_i < K_j$ for all $j \neq i$. In the asymmetric game, the mixed strategy equilibrium probability of

Number of other actors choosing C

Actor i's choice	0	1	2	...	$n-1$
Sanction (C)	$U_i - K_i$	$U_i - K_i$	$U_i - K_i$	$U_i - K_i$	$U_i - K_i$
Not sanction (D)	0	U_i	U_i	U_i	U_i

Fig 1. The volunteer's dilemma (VOD) game from actor i's perspective. In the VOD, actor i chooses between sanctioning (C) and declining to sanction (D). If i chooses C, i incurs a cost K_i but earns U_i as the collective good is produced. If i chooses D, i's payoff depends on what the other actors do. If all other actors choose D, the collective good is not produced and no one earns anything. If at least one of the other actors chooses C, the collective good is produced and i earns U_i. Since in a VOD $U_i > K_i > 0$, choosing D is not a dominant strategy, yet all actors prefer another actor to choose C while they choose D.

cooperation is: $p_i^* = 1 - \frac{U_i}{K_i} \left[\prod_{j=1}^{n} \frac{K_j}{U_j} \right]^{\left(\frac{1}{n-1} \right)}$. Paradoxically, this equilibrium implies that the strong actor chooses C with the lowest probability [35]. The mixed strategy equilibrium is, however, not evolutionary stable. The evolutionary stable equilibrium is the strategy profile in which the strong player cooperates while all other players defect, irrespective of group size. This asymmetric, pure strategy equilibrium is also selected by the Harsanyi-Selten rationality theory [36].

Axioms of strategic rationality imply the selection of the mixed Nash equilibrium in the symmetric VOD. In contrast, in the asymmetric game a unique asymmetric equilibrium is selected where the "strongest" player (the actor with the largest difference between U and K) cooperates and all other, "weaker" actors defect [35,36]. Experiments on strategic human behaviour support these predictions [37], and analyses of replicator dynamics demonstrate that the strategy profile "strongest actor cooperates / weaker actors defect" constitutes an evolutionary stable equilibrium [38] (see the online Supporting Information, S1 Additional study material and analyses).

The game theoretic analysis of the symmetric and asymmetric VOD allows for an interesting and important insight. At least in theory, asymmetry solves the VOD because the strongest actor will always provide the collective good–independent of group size. As a consequence, diffusion of responsibility is only to be expected in the symmetric VOD, whereas it will be absent if the VOD is asymmetric. Note that this is different from solutions that rely on an established leader or status hierarchies [39,40,41,42]. Appointing a leader to produce the public good alone solves the VOD exogenously and presupposes the existence of a legitimate authority. Here we argue that arbitrary, observable differences in individuals' costs of and/or benefits from producing the collective good can solve the VOD endogenously because group members tacitly coordinate on the strongest member to volunteer [15]. Note, moreover, that this mechanism may only be effective in solving step-level collective goods problems, in which the full contribution of a specific number of group members is required. In experiments with linear collective goods games, arbitrary differences in subjects' endowments or payoffs have been shown to have counterproductive effects on the efficient production of these collective goods [43,44,45,46].

In what follows, we experimentally test the hypothesis that the diffusion of responsibility effect decreases as a consequence of the switch from the symmetric to the asymmetric VOD. Cramer et al. [47] were the first to observe this effect in a field experiment, in which they varied group size ($n = 1$ vs $n = 2$) and competence of the bystander (trained nurse vs student). They found that, in the presence of an observer, trained nurses were as likely to help when hearing someone falling from a ladder as subjects in the absence of an observer, whereas students'

likelihood of helping was lower. Here we study this effect by means of a computerized laboratory experiment. We use the game theoretic model of the symmetric and asymmetric VOD to derive predictions about subjects' propensity to enforce a social norm. How social norms are enforced through sanctions has been of long-lasting interest to social scientists in general and sociologists in particular [48,49,50,51,52,53]. We argue and show that already relatively small differences in subjects' sanctioning costs can substantially attenuate diffusion of responsibility in norm enforcement if only one person is necessary and sufficient to enforce the social norm (see also [54]).

Methods

Experimental game

We use the "stealing game" with a sanctioning option to emulate a situation in which a norm violation can be negatively sanctioned and this sanction can be subject to diffusion of responsibility [37,55]. The stealing game is a simple representation of a situation in which one actor has the possibility to benefit from breaking a social norm at a cost for others. For example, someone who enjoys listening to loud music in public transport benefits from breaking a social norm but, at the same time, "steals" utility from others by producing negative externalities [54,56]. In this situation the intervention of only one person is necessary to stop the norm breaker. Accordingly, in the stealing game an actor X has the opportunity to steal a certain amount (U) from each of n other group members. If X steals, he or she gains nU, but each of the n "victims" has the possibility to sanction the thief thereby reclaiming the stolen money. A sanction is a costly cooperative action (with costs K) and reclaims the stolen money (U) for all n victims, irrespective of whether they had sanctioned or not. We further introduce a penalty condition in which a thief is not only obliged to pay back the stolen money but also incurs a penalty, if at least one of the n victims sanctions. With $U > K > 0$, the n victims face a volunteer's dilemma, irrespective of whether the thief can incur a penalty or not [8,37].

Experimental design

In our experiment, we vary the number of potential victims ($n = 2$ vs $n = 5$) and the type of VOD game (symmetric vs asymmetric) in a 2×2 between-subject factorial design. Every subject makes decisions in 30 consecutive rounds of the stealing game with sanctioning option; first 15 rounds without and then 15 with a penalty threat for the thief. In each round, the role of X, the role of potential victims, and the role of strong and weak subjects in the asymmetric VOD are assigned randomly. We assign subjects' roles randomly rather than having them assume the same role in every round for two reasons. First, experiencing the game in different roles allows subjects to better appreciate the strategic nature of the decision situation. Second, aggregate behaviour in particular roles is less dependent on the idiosyncrasies of the subjects assuming these roles. Moreover, we use stranger matching, i.e. after every round, groups are disbanded and randomly formed anew.

In every round, subject X first decides whether to steal 50 monetary units (MU) from each of the other n group members. If X steals, each victim decides independently whether or not to sanction to reclaim the stolen amount for the entire group ($U = 50$). Sanctioning costs are $K = 25$ for subjects in the symmetric VOD and for the strong subject in the asymmetric VOD, while sanctioning costs are higher ($K = 35$) for the weak subjects in the asymmetric VOD. In the second 15 rounds, a thief also incurs a penalty of 60 MU, if at least one of the victims sanctions. Hence, the second part of our experiment allows us to assess the impact of negative monetary sanctions on subjects' behaviour in the stealing game. Previous research demonstrates that negative monetary sanctions can have a substantial, positive effect on cooperation

in linear collective goods games [57,58,59]. Correspondingly, in our case, we expect the penalty to have a strong negative effect on the stealing rate. However, a low stealing rate will distort the conditions under which we can test our hypotheses on the diffusion of responsibility in norm enforcement. Therefore, we introduce the penalty only in the second part, i.e. after the 15 rounds without an additional penalty (also see [37]). Table 1 summarizes the parametrization of our experimental game and our experimental design.

Our design does not allow us to disentangle the effect of diffusion of responsibility on first-order cooperation (i.e. thieves' propensity to steal) and the effect of higher expected payoffs from stealing in larger groups. This is because U is constant across conditions whereas the amount that can be stolen increases with group size (nU). In other words, thieves' incentives to steal increase both because of the larger amount that can be stolen *and* the lower probability of a sanction in larger groups. We chose to keep U rather than nU constant across group size because the main aim of this study is to test the effect group heterogeneity has on the diffusion of responsibility in social norm enforcement. The effect of changing sanctioning probabilities on first-order cooperation has been corroborated elsewhere [15,37]. We therefore refrain from making predictions about the effect of experimental conditions on first-order cooperation (i.e. stealing). However, in the Results section we also report our findings regarding the stealing rate.

Predictions and hypotheses

Based on our parametrization of the VOD specified in Table 1, we can derive actors' probabilities of enforcing the social norm (i.e. sanction), the probabilities that at least one group member will sanction and the social norm will be enforced, and the probabilities that exactly one group member will sanction and the social norm will be enforced efficiently. The point predictions are listed in the upper half of Table 2; the formal derivation of these predictions is explained in all detail elsewhere (see [15,37]).

In the symmetric VOD, actors' sanctioning probabilities are predicted to be the same for all actors in a group and are denoted with p_i^*. In groups of $n = 2$, $p_i^* = 0.500$, and in groups of $n = 5$, $p_i^* = 0.159$. In other words, an actor's probability to sanction decreases with group size, which corresponds to the diffusion of responsibility effect. The diffusion of responsibility effect is also predicted at the group level. The probability that at least one actor sanctions is $p^* = 0.750$ in groups of $n = 2$, and it is $p^* = 0.580$ in groups of $n = 5$. The diffusion of responsibility will also be reflected in the probability that exactly one actor sanctions and enforces the social norm efficiently, which is $p_e^* = 0.500$ in groups of $n = 2$, and $p_e^* = 0.398$ in groups of $n = 5$.

In contrast, in the asymmetric VOD, the "weak" actors are predicted to sanction with probability $p_w^* = 0$ and the strong actors sanction with probability $p_s^* = 1$, irrespective of group size. Hence, no diffusion of responsibility is predicted to occur in the asymmetric VOD,

Table 1. Experimental games and experimental design.

	symmetric		asymmetric	
	$n = 2$ **$U = 50$ MU** **$K = 25$ MU**	**$n = 5$** **$U = 50$ MU** **$K = 25$ MU**	**$n = 2$** **$U = 50$ MU** **$K_s = 25$ MU** **$K_w = 35$ MU**	**$n = 5$** **$U = 50$ MU** **$K_s = 25$ MU** **$K_w = 35$ MU**
part 1: rounds 1–15	no penalty $P = 0$ MU	no penalty $P = 0$ MU	no penalty $P = 0$ MU	no penalty $P = 0$ MU
part 2: rounds 16–30	Penalty $P = 60$ MU	Penalty $P = 60$ MU	Penalty $P = 60$ MU	Penalty $P = 60$ MU

n: group size; U: Benefit from cooperation; K: cost of cooperation; K_s: cost of cooperation for strong player; K_w: cost of cooperation for weak player; P: penalty.

Table 2. Predicted sanctioning probabilities.

	symmetric		asymmetric	
	$n = 2$	$n = 5$	$n = 2$	$n = 5$
p_i^*	0.500	0.159	n/a	n/a
p_s^*	n/a	n/a	1	1
p_w^*	n/a	n/a	0	0
p^*	0.750	0.580	1	1
p_e^*	0.500	0.398	1	1

n: group size; p_i^*: equilibrium sanctioning probability; p_s^*: equilibrium sanctioning probability of strong player; p_w^*: equilibrium sanctioning probability of weak player; p^*: probability that at least one sanction will occur in equilibrium; p_e^*: probability that exactly one sanction will occur in equilibrium.

https://doi.org/10.1371/journal.pone.0208129.t002

neither at the individual nor at the group level ($p^* = 1$), and the social norm is always enforced by the strong actor alone ($p_e^* = 1$). Based on these point predictions, we can state the following hypotheses:

Symmetric VOD: n = 2 vs n = 5

H1.1: In the symmetric VOD, subjects will less often sanction in groups of $n = 5$ than in groups of $n = 2$.

H1.2: In the symmetric VOD, the social norm will less often be enforced in groups of $n = 5$ than in groups of $n = 2$.

H1.3: In the symmetric VOD, the social norm will less often be enforced *efficiently* in groups of $n = 5$ than in groups of $n = 2$.

Asymmetric VOD: weak subject vs strong subject

H2: In the asymmetric VOD, "weak" subjects will less often sanction than "strong" subjects.

Asymmetric VOD: n = 2 vs n = 5

H3.1: In the asymmetric VOD, "strong" subjects will as often sanction in groups of $n = 5$ as in groups of $n = 2$.

H3.2: In the asymmetric VOD, the social norm will as often be enforced in groups of $n = 5$ as in groups of $n = 2$.

H3.3: In the asymmetric VOD, the social norm will as often be enforced *efficiently* in groups of $n = 5$ as in groups of $n = 2$.

Symmetric VOD vs asymmetric VOD

H4.1: In the symmetric VOD, (weak) subjects will less often sanction than "strong" subjects in the asymmetric VOD.

H4.2: In the symmetric VOD, the social norm will less often be enforced than in the asymmetric VOD.

H4.3: In the symmetric VOD, the social norm will less often be enforced *efficiently* than in the asymmetric VOD.

Experimental procedures

In total, 252 subjects participated in our computerized laboratory experiment. The experiment comprised seven sessions and 36 subjects participated in each session. Because of the random matching protocol and the large groups size in one of the treatments, we varied group size between sessions. The treatment with five victims ($N = 5 + 1$) was conducted in sessions 1, 3, 5 and 7, and the treatment with two victims ($N = 2 + 1$) was conducted in sessions 2, 4 and 6. In sessions testing the treatment with two victims, subjects were randomly assigned to the symmetric VOD (subjects 1 through 18) or asymmetric VOD (subjects 19 through 36). In the treatment with five victims the type of VOD was varied across sessions.

The experiment was conducted in the Decision Science Laboratory at ETH Zurich (DeSciL). Subjects were students from the University of Zurich and ETH Zurich, 59.5% were female and they were 22.7 years old on average (sd = 3.00). Upon arrival in the lab, subjects received condition-specific instructions on paper. The instructions that were given to subjects in one of the experimental conditions (asymmetric VOD with $n = 5$) are provided in the online Supporting Information (S1 Additional study material and analyses). Instructions explained the decision situations step by step and contained shots of the actual decision screens. Moreover, subjects learned that their decisions were anonymous, that their payments would correspond to the sum they earned in each round and that payments would be administered by a person not involved in the implementation of the experiment. After reading the instructions, subjects took a quiz with questions about the decision situations. Questions for which at least one wrong answer was given were read out loud and the correct answer was explained to all subjects at the same time. Then, the experiment started. A session lasted for about 1h and subjects earned CHF 38 (incl. CHF 10 show-up fee) on average ($\approx$ USD 43.4). After the experiment, subjects filled in a questionnaire and could leave the lab to get their payment in private. The experiment was programmed and conducted with the software z-Tree [60].

Data analysis strategy

All result figures and test statistics are based on regression model estimations. We only use logistic regression models because all our target variables are binary [61]. Statistical significance is set at the 5% level (i.e. $\alpha = 0.05$) for two-sided tests and we account for the repeated measures obtained on the same subject by estimating cluster-robust standard errors [62,63]. We use Stata's margins command to calculate proportions from saturated logit models and test the statistical significance of the differences between proportions using Wald tests of linear hypotheses. The data and Stata 14 script producing the results and figures reported in this paper are available as online Supporting Information (S1 Data, S1 Stata do-file).

Ethics statement

This research was approved by the Ethics Committee of ETH Zurich. All the research was performed in the Decision Science Laboratory(DeSciL) at ETH Zurich, Haldeneggsteig 4, CH-8092 Zurich, Switzerland. The review board of DeSciL is called DeSciL Review Board, and its members are listed on the DeSciL website (https://www.descil.ethz.ch/people). Our experiment was conducted in accordance with DeSciL Operational Rules, which are approved by the review board and published on the DeSciL website (https://www.descil.ethz.ch/lab/researchers). All participants in our experiment were recruited from the subject pool maintained by the University Registration Center for Study Participants (UAST) of the University of Zurich and ETH Zurich. Every person who has signed up to this subject pool also gave his or her consent by agreeing to the terms and conditions of UAST via the online registration tool. Consent is required for registration and is given by selecting the checkbox labelled "I have read the terms and conditions and I agree with them." These terms and conditions are published on the UAST website (https://www.uast.uzh.ch/register). Terms and conditions inform about the procedure of PC decision experiments in general, the anonymity of the data, and the usage of the anonymized data for the purpose of scientific studies.

Results

Fig 2 shows stealing rates (i.e. social norm violations) across experimental conditions. Across all conditions there is a substantial and significant penalty effect. That is, the stealing rates are considerably lower in the last 15 rounds as compared to the first 15 rounds, and the overall
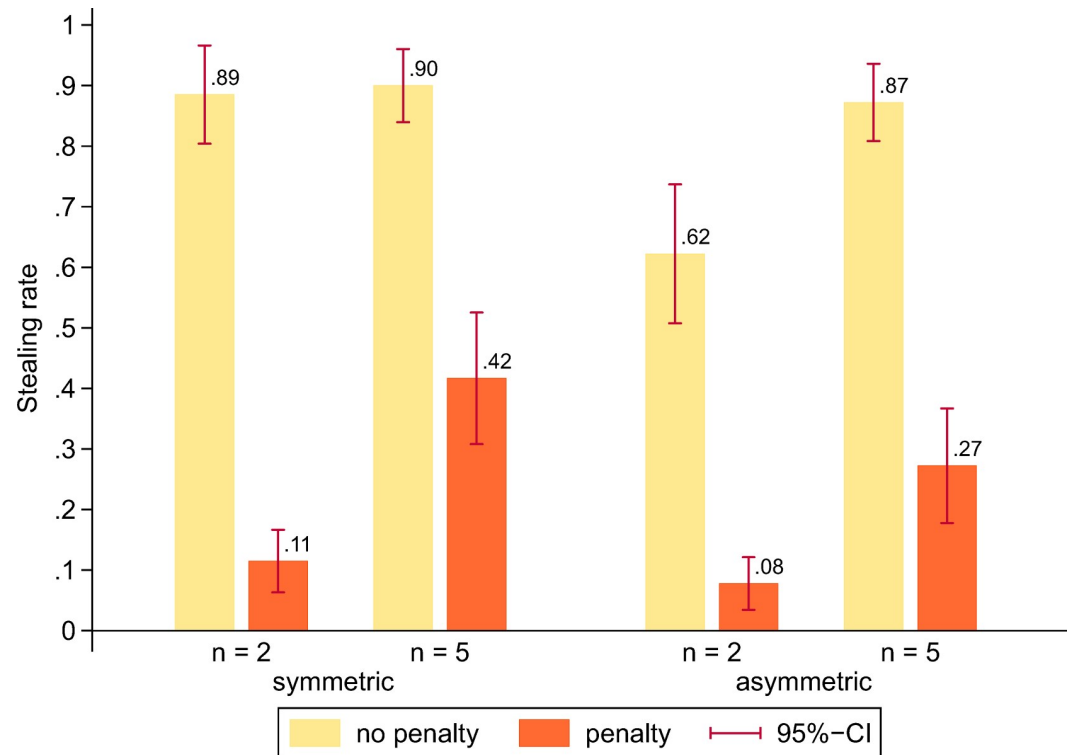
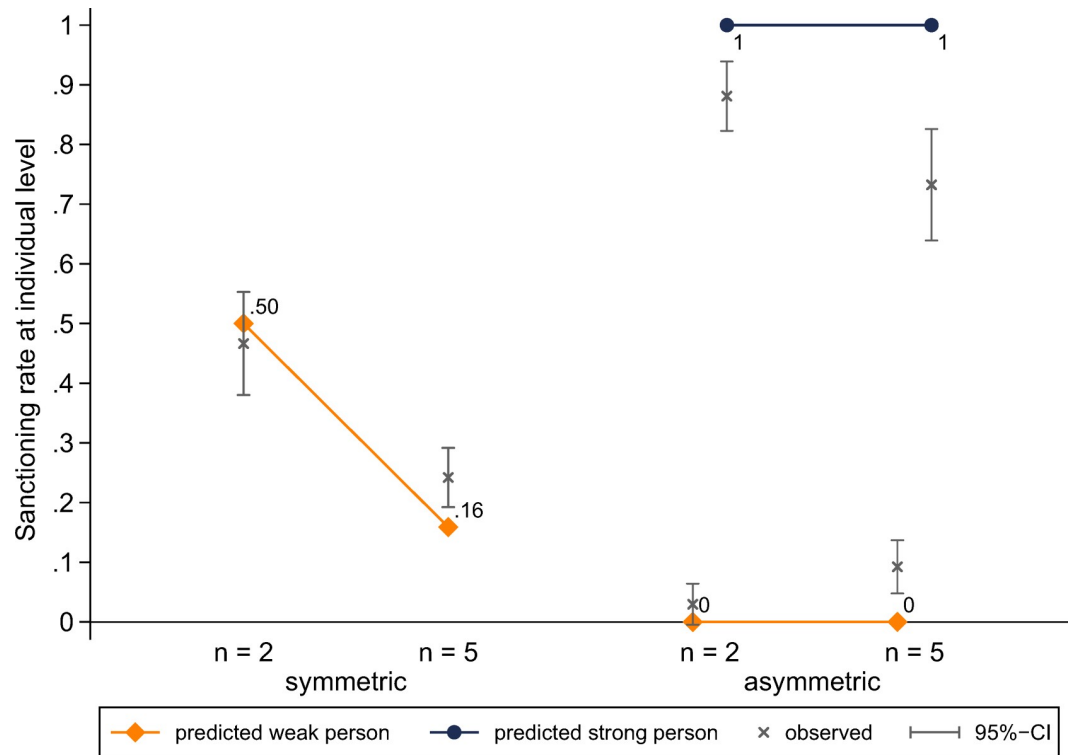**Fig 2. Stealing rates across experimental conditions.**

reduction is significantly larger in groups of $n = 2$ than in groups of $n = 5$ ($z = 3.08$, $p = 0.002$). In groups of $n = 2$ there is no significant difference in the stealing rate between the symmetric and the asymmetric condition with penalty ($z = 1.46$, $p = 0.144$). In groups of $n = 5$, the stealing rate is significantly lower in the asymmetric condition than the symmetric condition ($z = 2.92$, $p = 0.004$). Without penalty (in the first 15 rounds), only the stealing rate in the asymmetric $n = 2$ condition is significantly lower than the stealing rates in the other three conditions (see Fig 2). Recall that our experimental design does not allow us to disentangle the effects of diffusion of responsibility and expected payoffs of stealing on stealing rates because both factors increase the incentive to steal as groups become larger (see Experimental design). We therefore refrain from stating and testing hypotheses relating experimental conditions to stealing behavior.

The penalty effect leads to less stealing and thus a much smaller number of observations at the sanctioning stage (Fig 2). We therefore restrict our reporting of main results to the first 15 rounds, without a penalty threat for the thief. The analysis of the second 15 rounds, with a potential penalty for the thief, yields similar results and is reported in the online Supporting Information (S1 Additional study material and analyses). In the following we first report results pertaining to individual level behaviour (i.e. sanctioning rate), then the results regarding the rate of social norm enforcement (i.e. collective good production) and lastly the results regarding the rate of *efficient* social norm enforcement.

Testing hypotheses

Fig 3 shows predicted and observed individual sanctioning rates across experimental conditions and subject roles. In line with hypothesis H1.1, we observe a strong diffusion of
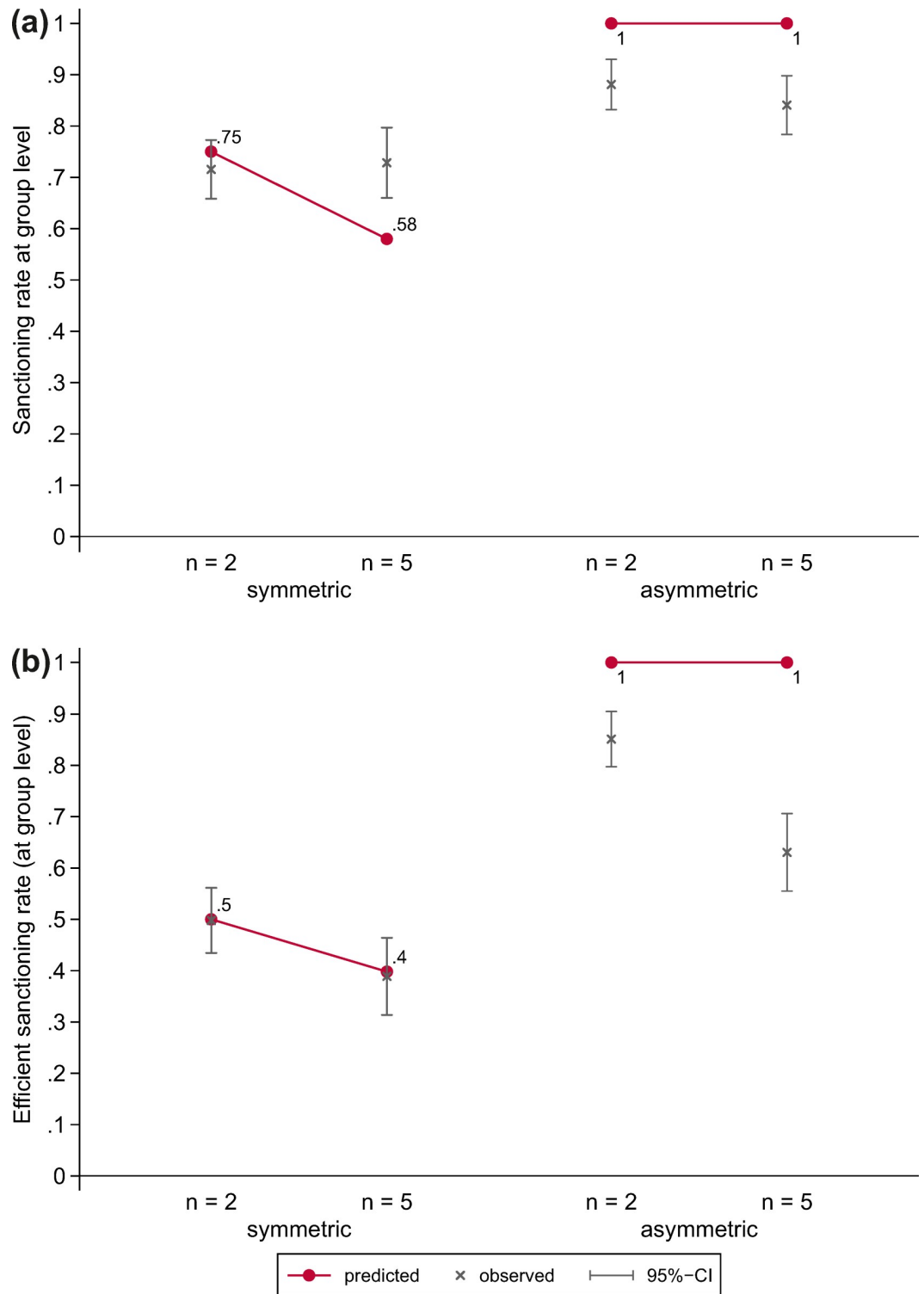
Fig 3. Predicted and observed sanctioning rates at the individual level across experimental conditions and subject roles.

responsibility effect in the symmetric conditions. The decrease in the individual sanctioning rate from 0.47 to 0.24 if group size increases from two to five is statistically significant ($z = 4.42$, $p < 0.001$). Moreover, we observe a substantial disparity of subjects' behaviour between the symmetric and asymmetric VOD. In the asymmetric VOD, in line with hypotheses H2, strong subjects sanction and weak subjects refrain from sanctioning almost all the time. These results are also in line with hypothesis H4.1, according to which weak subjects in the asymmetric VOD will sanction less than (weak) subjects in the symmetric VOD ($n = 2$: $z = 9.20$, $p < 0.001$; $n = 5$: $z = 4.39$, $p < 0.001$). Subjects' behaviour is thus very much in accordance with the asymmetric equilibrium predictions [35,36,38]. As expected, the decrease in the individual sanctioning rate across group size is smaller in the asymmetric dilemma (from 0.88 to 0.73), but the difference is still significant ($z = 2.65$, $p = 0.008$). Moreover, the difference in differences is statistically insignificant ($z = 1.01$, $p = 0.315$). These result go clearly against our hypothesis H3.1. We will come back to this point shortly.

Fig 4A shows the predicted and observed sanctioning rates at the group level across experimental conditions. The sanctioning rate at the group level denotes the proportion of cases in which *at least one* victim sanctioned the thief. We find no support for hypothesis H1.2; there is no statistically significant decrease in the sanctioning rate when group size increases from $n = 2$ to $n = 5$ in the symmetric VOD ($z = -0.28$, $p = 0.777$). That is, unlike at the individual level, there is no evidence for a decrease of collective good production at the group level. At the same time, we cannot reject hypothesis H3.2; there is no significant difference between sanctioning rates in groups of size $n = 2$ and $n = 5$ in the asymmetric VOD ($z = 1.04$, $p = 0.296$). We find clear support for hypothesis H4.2; the social norm is less often enforced in the symmetric than in the asymmetric VOD both in groups of $n = 2$ ($z = -4.30$, $p < 0.001$) and in groups of $n = 5$ ($z = -2.47$, $p = 0.014$).

Fig 4. Predicted and observed sanctioning rates at the group level across experimental conditions. (a) Sanction by at least one person. (b) Sanction by exactly one person].

https://doi.org/10.1371/journal.pone.0208129.g004

The efficient sanctioning rate (both predicted and observed) presented in Fig 4B denotes the proportion of cases in which *exactly one* victim sanctioned the thief. In line with the diffusion of responsibility hypothesis (H1.3), the figure shows that group size has a negative effect on the efficient sanctioning rate in the symmetric VOD ($z = 2.17$, $p = 0.030$). However, a significant decrease in the efficient sanctioning rate with group size can also be observed in the asymmetric VOD ($z = 4.66$, $p < 0.001$), which leads us to reject hypothesis H3.3. Still, we find clear support for our hypothesis (H4.3) that the second order public good is less often produced efficiently in the symmetric VOD than in the asymmetric VOD both in groups of $n = 2$ ($z = -8.32$, $p < 0.001$) and in groups of $n = 5$ ($z = -4.45$, $p < 0.001$).

Out of our ten hypotheses, seven are supported and three are rejected by the empirical evidence. The three rejected hypotheses are on the diffusion of responsibility effect (three out of six), one in the symmetric VOD predicting diffusion of responsibility to occur with an increase in group size (H1.2) and two in the asymmetric VOD predicting the absence of diffusion of responsibility (H3.1 and H3.3). We have good reason to assume that this is not a stable result. In what follows, we formulate our conjectures and perform exploratory analyses using the same data. However, the findings resulting from this exploratory analysis can only be interpreted as first evidence and must be verified in further experiments.

Exploratory analyses

Our conjecture is that in the asymmetric VOD subjects take longer to tacitly coordinate on the strong subject to sanction the thief alone if groups are larger. In the asymmetric VOD with two players, the difference in payoffs becomes immediately apparent; in the asymmetric VOD with five players, it might take a while for subjects to realize that one of them has lower costs for sanctioning (despite the fact that this is made clear in the instructions). Hence, over time, subjects will learn to appreciate the strategic nature of the decision situation and tacitly coordinate on the strong subject to sanction alone. Although we did not *a priori* hypothesize a "learning effect" (in fact an interaction between learning and group size in the asymmetric VOD), we can control for time in a more detailed, exploratory analysis.

We divide the 15 rounds without penalty (see Table 1) in blocks of five rounds (i.e. rounds 1 to 5, 6 to 10, and 11 to 15) and perform additional analyses to test our conjecture. The division in three blocks of five rounds is a natural choice because with fewer blocks we may not be able to capture the time trend implicit in our conjecture and with more blocks (e.g. five blocks of three rounds) the low number of cases in the last block will decrease the power of the statistical tests of our conjecture. We did not perform any additional analysis based on other divisions of the fifteen rounds.

In the same way as Fig 3, Fig 5 shows predicted and observed individual sanctioning rates across experimental conditions and subject roles. But unlike in Fig 3, the observed rates are also shown over time. In support of our conjecture, Fig 5 shows that the sanctioning rate of the strong subject increases over time in the $n = 5$ condition. As a result, in the last five rounds (i.e. rounds 11 to 15), the group size effect in the asymmetric VOD is not significant anymore ($z = 0.77$, $p = 0.441$). Hence, subjects seem to take longer to adjust their behaviour in the more complex decision task with five group members. Now we also find that, in the last five rounds, the difference in differences between the symmetric and asymmetric VOD is statistically significant (0.90–0.84 vs 0.52–0.23; $z = 2.50$, $p = 0.012$). That is, if group size increases from two to five, the decrease in the strong player's sanctioning rate is significantly smaller than the decrease in the individual sanctioning rate in the symmetric VOD.

Fig 6A shows the predicted and observed group-level sanctioning rates (where the sanction was performed by at least one subject) across experimental conditions and time. Although the
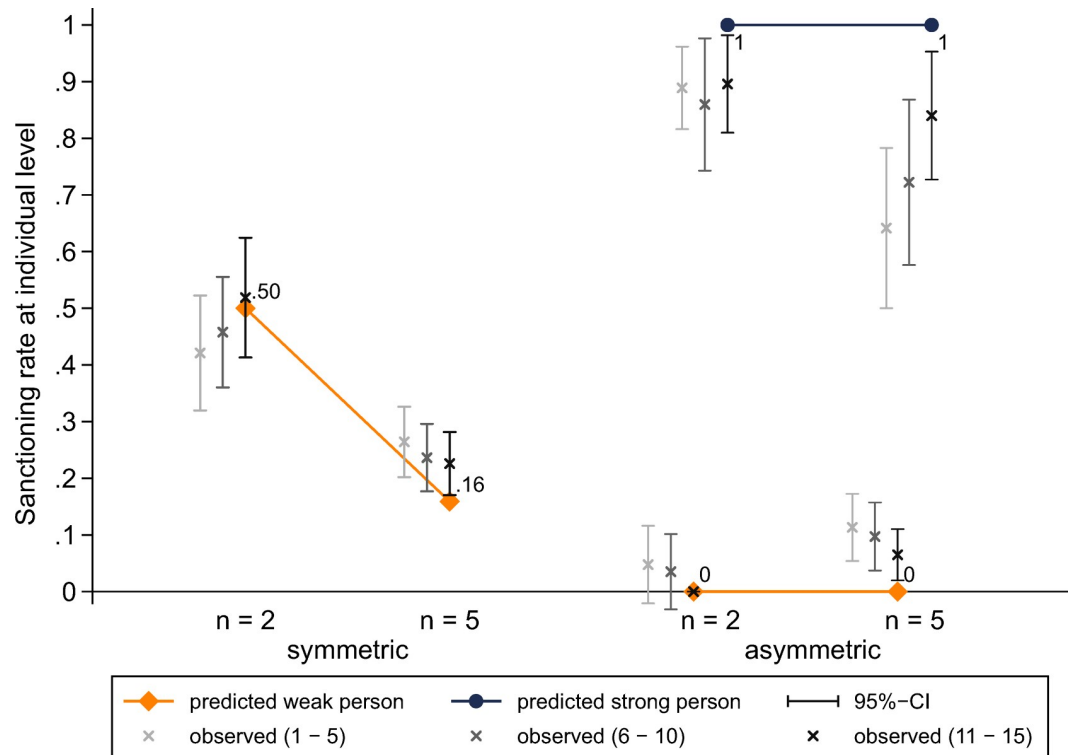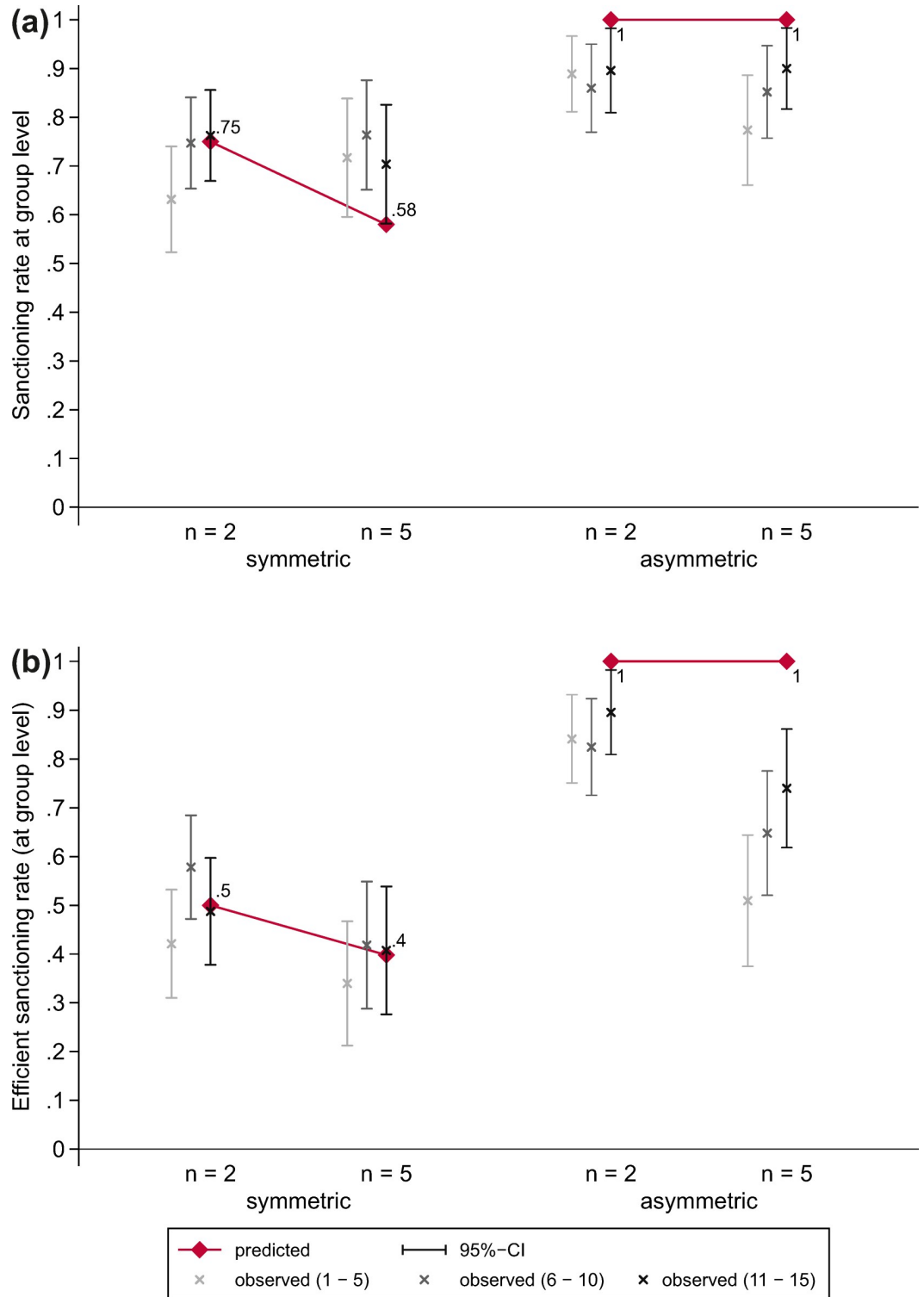
Fig 5. Predicted and observed sanctioning rates at the individual level across experimental conditions, subject roles and time.

diffusion of responsibility effect in the symmetric VOD remains absent (i.e. H1.2 remains unsupported), the figure shows that in the asymmetric VOD with $n = 5$ subjects' behavior converges to the predicted sanctioning rate. Fig 6B shows the predicted and observed efficient sanctioning rates (where the sanction was performed by exactly one subject) across experimental conditions and time. Although there is a statistically significant increase in the efficient sanctioning rate in the asymmetric VOD with $n = 5$ from the first five to the last five rounds ($z = 2.49$, $p = 0.013$), the sanctioning rate in the last five rounds remains significantly lower than the rate in the asymmetric VOD with $n = 2$ ($z = 2.05$, $p = 0.041$). All in all our exploratory analyses corroborate our conjecture that subjects need more time to reach the sanctioning rates predicted in the asymmetric VOD if groups are larger.

Discussion

Five decades ago, Darley and Latané [19] showed that an emergency situation may end tragically despite the presence of potential helpers if each of them expects someone else to intervene. The authors called this mechanism "diffusion of responsibility." Their recognizing the mechanism and corroborating it empirically was an important step towards a better understanding of the dynamics and outcomes of emergency situations. Although the observation of diffusion of responsibility in different types of volunteering situations is very robust [20,30], there are also exceptions. Here we argue that observable asymmetries of the bystanders in terms of their costs of and/or benefits from intervening may account for these exceptions (see also [47,64,65]). To test this conjecture, we model the volunteering situation as a symmetric and asymmetric volunteer's dilemma game (VOD), derive hypotheses about the moderating

Fig 6. Predicted and observed sanctioning rates at the group level across experimental conditions and time. (a) Sanction by at least one person. (b) Sanction by exactly one person.

effect of asymmetry on the diffusion of responsibility in social norm enforcement, and test our predictions in a computerized lab experiment.

In our experiment, we use the stealing game with a sanctioning option to emulate a situation in which a norm violation can be negatively sanctioned. That is, in a group of subjects, one randomly chosen subject can decide to steal money from the other group members (i.e. violate a social norm), who then face a VOD as only one of them is required to reclaim the stolen amount for the entire group (i.e. enforce the social norm). We vary group size and symmetry of the VOD in terms of the costs a subject incurs from sanctioning the thief. Our results show a clear diffusion of responsibility effect in the symmetric VOD, in which all group members have the same costs of sanctioning the thief. In the asymmetric VOD, diffusion of responsibility is largely diminished but only after subject have played the game for some time. By and large, these results are also borne out at the group level. In particular, heterogeneous groups become more effective in enforcing social norms as they manage to tacitly coordinate on the strongest subject to sanction the norm breaker alone. Our findings support the proposition that even relatively small asymmetries in observable sanctioning costs facilitate bystanders' tacit coordination on the "strongest" individual to negatively sanction norm breakers. In other words, our results show how asymmetry can "break" the diffusion of responsibility in social norm enforcement and help to overcome the second-order free-rider problem [53,66,67].

Our findings complement results from field experiments in social psychology in which the perception of competent bystanders inhibits cooperation of other group members [55,68]. There is also supporting evidence from research on coordination behaviour in non-human primates and small children [69,70]. In groups of chimpanzees the group member with highest rank takes solitary action in a volunteer's dilemma [18]. Because the chimpanzee's rank is negatively correlated with the costs of cooperation, this experimental observation is in line with our results on the asymmetric VOD. Asymmetry can also establish a focal point for individuals to tacitly coordinate on a unique equilibrium in a situation with multiple equilibria [15,71,72]. Such an ability to tacitly coordinate on a focal equilibrium seems to emerge in children as from the age of five [73].

In how far asymmetry in costs from intervening provides an incentive for, and in how far it merely establishes a focal point facilitating coordination on the "strongest" individual to volunteer in the asymmetric VOD is an open question to be addressed in future research. Future research should also reconsider the effect of individual heterogeneity on collective action and cooperation in groups. Although predicted by theory [35,36], it is an empirical question whether introducing payoff differences in weak subjects as well (while keeping one identifiable strong subject) would produce similar results as reported above. Our conjecture is that a gradual distinction of subjects strength in terms of payoffs (costs and or benefits of intervening) would also facilitate tacit coordination on the strongest subset of $k > 1$ subjects that is required to produce a collective good for a group of size $n > k$. This would extend the applicability of our findings beyond the volunteer's dilemma, for which $k = 1$.

Supporting information

S1 Additional study material and analyses.
(PDF)

S1 Data.
(DTA)

S1 Stata do file.
(DO)

Acknowledgments

Author Contributions

Conceptualization: Wojtek Przepiorka, Andreas Diekmann.

Formal analysis: Wojtek Przepiorka, Andreas Diekmann.

Investigation: Wojtek Przepiorka.

Methodology: Wojtek Przepiorka, Andreas Diekmann.

Software: Wojtek Przepiorka.

Validation: Wojtek Przepiorka.

Visualization: Wojtek Przepiorka.

Writing – original draft: Wojtek Przepiorka, Andreas Diekmann.

Writing – review & editing: Wojtek Przepiorka, Andreas Diekmann.

References

1. Archetti M, Scheuring I. Coexistence of cooperation and defection in public goods games. Evolution. 2011; 65(4):1140–8. https://doi.org/10.1111/j.1558-5646.2010.01185.x PMID: 21062277

2. Capraro V, Barcelo H. Group Size Effect on Cooperation in One-Shot Social Dilemmas II: Curvilinear Effect. PLOS ONE. 2015; 10(9):e0138744. https://doi.org/10.1371/journal.pone.0138744 PMID: 26376305

3. Oliver PE, Marwell G, Teixeira R. A Theory of the Critical Mass. I. Interdependence, Group Heterogeneity, and the Production of Collective Action. American Journal of Sociology. 1985; 91(3):522–56.

4. Olson M. The Logic of Collective Action: Public Goods and the Theory of Groups. Cambridge (MA): Harvard University Press; 1971 [1965].

5. Rapoport A, Suleiman R. Incremental Contribution in Step-Level Public Goods Games with Asymmetric Players. Organizational Behavior and Human Decision Processes. 1993; 55(2):171–94.

6. Archetti M, Scheuring I. Review: Game theory of public goods in one-shot social dilemmas without assortment. Journal of Theoretical Biology. 2012; 299:9–20. https://doi.org/10.1016/j.jtbi.2011.06.018 PMID: 21723299

7. Palfrey TR, Rosenthal H. Participation and the Provision of Discrete Public Goods: A Strategic Analysis. Journal of Public Economics. 1984; 24(2):171–93.

8. Raihani NJ, Bshary R. The Evolution of Punishment in n-Player Public Goods Games: A Volunteer's Dilemma. Evolution. 2011; 65(10):2725–8. https://doi.org/10.1111/j.1558-5646.2011.01383.x PMID: 21967415

9. van de Kragt AJC, Orbell JM, Dawes RM. The Minimal Contributing Set as a Solution to Public Goods Problems. American Political Science Review. 1983; 77(1):112–22.

10. Diekmann A. Volunteer's Dilemma. Journal of Conflict Resolution. 1985; 29(4):605–10.

11. Goeree JK, Holt CA, Smith AM. An experimental examination of the volunteer's dilemma. Games and Economic Behavior. 2017; 102:303–15.

12. Archetti M. A Strategy to increase cooperation in the volunteer's dilemma: Reducing vigilance improves alarm calls. Evolution. 2011; 65(3):885–92. https://doi.org/10.1111/j.1558-5646.2010.01176.x PMID: 21091467

13. Lee S, Levin D, Gopalakrishnan V, Bhattacharjee B. Backbone Construction in Selfish Wireless Networks. ACM SIGMETRICS Performance Evaluation Review. 2007; 35(1):121–36.

14. Naserian M, Tepe K. Game theoretic approach in routing protocol for wireless ad hoc networks. Ad Hoc Networks. 2009; 7(3):569–78.

15. Diekmann A, Przepiorka W. Punitive preferences, monetary incentives and tacit coordination in the punishment of defectors promote cooperation in humans. Scientific Reports. 2015; 5:10321. https://doi.org/10.1038/srep10321 PMID: 25988875

16. Diekmann A, Przepiorka W. "Take One for the Team!" Individual heterogeneity and the emergence of latent norms in a volunteer's dilemma. Social Forces. 2016; 93(3):1309–33.

17. Weesie J, Franzen A. Cost Sharing in a Volunteer's Dilemma. Journal of Conflict Resolution. 1998; 42 (5):600–18.

18. Schneider A-C, Melis AP, Tomasello M. How chimpanzees solve collective action problems. Proceedings of the Royal Society B. 2012; 279:4946–54. https://doi.org/10.1098/rspb.2012.1948 PMID: 23075841

19. Darley JM, Latané B. Bystander Intervention in Emergencies: Diffusion of Responsibility. Journal of Personality and Social Psychology. 1968; 8(4):377–83. PMID: 5645600

20. Fischer P, Krueger JI, Greitemeyer T, Vogrincic C, Kastenmüller A, Frey D, et al. The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. Psychological Bulletin. 2011; 137(4):517–37. https://doi.org/10.1037/a0023304 PMID: 21534650

21. Kollock P. Social Dilemmas: The Anatomy of Cooperation. Annual Review of Sociology. 1998; 24:183–214.

22. Raub W, Buskens V, Corten R. Social Dilemmas and Cooperation. In: Braun N, Saam NJ, editors. Handbuch Modellbildung und Simulation in den Sozialwissenschaften. Wiesbaden: Springer VS; 2015. p. 597–626.

23. van Lange PAM, Balliet DP, Parks CD, van Vugt M. Social Dilemmas: Understanding Human Cooperation. Oxford: Oxford University Press; 2014.

24. Barcelo H, Capraro V. Group size effect on cooperation in one-shot social dilemmas. Scientific Reports. 2015; 5:7937. https://doi.org/10.1038/srep07937 PMID: 25605124

25. Bonacich P, Shure GH, Kahan JP, Meeker RJ. Cooperation and Group Size in the N-Person Prisoners' Dilemma. Journal of Conflict Resolution. 1976; 20(4):687–706.

26. Franzen A. Group-Size and One-Shot Collective Action. Rationality and Society. 1995; 7(2):183–200.

27. Ledyard JO. Public Goods: A Survey of Experimental Research. In: Kagel JH, Roth AE, editors. The Handbook of Experimental Economics. Princeton (NJ): Princeton University Press; 1995. p. 111–94.

28. Marwell G, Ames RE. Experiments on the Provision of Public-Goods. I. Resources, Interest, Group-Size, and the Free-Rider Problem. American Journal of Sociology. 1979; 84(6):1335–60.

29. Latané B, Darley JM. The Unresponsive Bystander: Why doesn't he help? New York: Appleton-Century-Crofts; 1970.

30. Latané B, Nida S. Ten years of research on group size and helping. Psychological Bulletin. 1981; 89 (2):308–24.

31. Latané B. The psychology of social impact. American Psychologist. 1981; 36(4):343–56.

32. Petty RE, Harkins SG, Williams KD, Latane B. The Effects of Group Size on Cognitive Effort and Evaluation. Personality and Social Psychology Bulletin. 1977; 3(4):579–82.

33. Piliavin IM, Piliavin JA, Rodin J. Costs, diffusion, and the stigmatized victim. Journal of Personality and Social Psychology. 1975; 32(3):429–38.

34. Archetti M. The volunteer's dilemma and the optimal size of a social group. Journal of Theoretical Biology. 2009; 261(3):475–80. https://doi.org/10.1016/j.jtbi.2009.08.018 PMID: 19703470

35. Diekmann A. Cooperation in an Asymmetric Volunteer's Dilemma Game: Theory and Experimental Evidence. International Journal of Game Theory. 1993; 22(1):75–85.

36. Harsanyi JC, Selten R. A General Theory of Equilibrium Selection in Games. Cambridge (MA): MIT Press; 1988.

37. Przepiorka W, Diekmann A. Individual heterogeneity and costly punishment: a volunteer's dilemma. Proceedings of the Royal Society B. 2013; 280(1759):20130247. https://doi.org/10.1098/rspb.2013.0247 PMID: 23536599

38. He J-Z, Wang R-W, Li Y-T. Evolutionary Stability in the Asymmetric Volunteer's Dilemma. PLOS ONE. 2014; 9(8):e103931. https://doi.org/10.1371/journal.pone.0103931 PMID: 25111781

39. Baldassarri D, Grossman G. Centralized sanctioning and legitimate authority promote cooperation in humans. Proceedings of the National Academy of Sciences USA. 2011; 108(27):11023–7.

40. Baumeister RF, Chesner SP, Senders PS, Tice DM. Who's in Charge Here? Group Leaders Do Lend Help in Emergencies. Personality and Social Psychology Bulletin. 1988; 14(1):17–22. https://doi.org/10.1177/0146167288141002 PMID: 30045447

41. O'Gorman R, Henrich J, Van Vugt M. Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation. Proceedings of the Royal Society B. 2009; 276 (1655):323–9. https://doi.org/10.1098/rspb.2008.1082 PMID: 18812292

42. Simpson B, Willer R, Ridgeway CL. Status Hierarchies and the Organization of Collective Action. Sociological Theory. 2012; 30(3):149–66.

43. Burton-Chellew MN, May RM, West SA. Combined inequality in wealth and risk leads to disaster in the climate change game. Climatic Change. 2013; 120(4):815–30.

44. Janssen MA, Anderies JM, Joshi SR. Coordination and cooperation in asymmetric commons dilemmas. Experimental Economics. 2011; 14(4):547–66.

45. Kölle F. Heterogeneity and cooperation: The role of capability and valuation on public goods provision. Journal of Economic Behavior & Organization. 2015; 109:120–34.

46. Reuben E, Riedl A. Enforcement of contribution norms in public good games with heterogeneous populations. Games and Economic Behavior. 2013; 77(1):122–37.

47. Cramer RE, Mcmaster MR, Bartell PA, Dragna M. Subject Competence and Minimization of the Bystander Effect. Journal of Applied Social Psychology. 1988; 18(13):1133–48.

48. Balafoutas L, Nikiforakis N. Norm enforcement in the city: A natural field experiment. European Economic Review. 2012; 56(8):1773–85.

49. Berger J, Hevenstone D. Norm enforcement in the city revisited: An international field experiment of altruistic punishment, norm maintenance, and broken windows. Rationality and Society. 2014; 28 (3):299–319.

50. Horne C. The Rewards of Punishment: A Relational Theory of Norm Enforcement. Stanford (CA): Stanford University Press; 2009.

51. Kitts JA. Collective Action, Rival Incentives, and the Emergence of Antisocial Norms. American Sociological Review. 2006; 71(2):235–59.

52. Piskorski MJ, Gorbatâi A. Testing Coleman's Social-Norm Enforcement Mechanism: Evidence from Wikipedia. American Journal of Sociology. 2017; 122(4):1183–222.

53. Yamagishi T. The provision of a sanctioning system as a public good. Journal of Personality and Social Psychology. 1986; 51(1):110–6.

54. Przepiorka W, Berger J. The sanctioning dilemma: A quasi-experiment on social norm enforcement in the train. European Sociological Review. 2016; 32(3):439–51.

55. Chekroun P, Brauer M. The bystander effect and social control behavior: the effect of the presence of others on people's reactions to norm violations. European Journal of Social Psychology. 2002; 32 (6):853–67.

56. Coleman JS. Foundations of Social Theory. Cambridge (MA): The Belknap Press of Harvard University Press; 1990.

57. Fehr E, Gächter S. Altruistic punishment in humans. Nature. 2002; 415(6868):137–40. https://doi.org/ 10.1038/415137a PMID: 11805825

58. Gürerk Ö, Irlenbusch B, Rockenbach B. The Competitive Advantage of Sanctioning Institutions. Science. 2006; 312(5770):108–11. https://doi.org/10.1126/science.1123633 PMID: 16601192

59. Herrmann B, Thöni C, Gächter S. Antisocial Punishment Across Societies. Science. 2008; 319 (5868):1362–7. https://doi.org/10.1126/science.1153808 PMID: 18323447

60. Fischbacher U. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics. 2007; 10(2):171–8.

61. Long JS. Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage; 1997.

62. Fitzmaurice GM, Laird NM, Ware JH. Applied Longitudinal Analysis. Hoboken (NJ): John Wiley & Sons; 2004.

63. Snijders TAB, Bosker RJ. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Los Angeles: Sage; 2012.

64. Brauer M, Chekroun P. The Relationship Between Perceived Violation of Social Norms and Social Control: Situational Factors Influencing the Reaction to Deviance. Journal of Applied Social Psychology. 2005; 35(7):1519–39.

65. Healy AJ, Pate JG. Cost asymmetry and incomplete information in a volunteer's dilemma experiment. Social Choice and Welfare. 2018; 51(3):465–91.

66. Heckathorn DD. Collective Action and the Second-Order Free-Rider Problem. Rationality and Society. 1989; 1(1):78–100.

67.   Oliver P. Rewards and Punishments as Selective Incentives for Collective Action: Theoretical Investigations. American Journal of Sociology. 1980; 85(6):1356–75.

68.   Bickman L. The effect of another bystander's ability to help on bystander intervention in an emergency. Journal of Experimental Social Psychology. 1971; 7(3):367–79.

69.   Brosnan SF, Beran MJ, Parrish AE, Price SA, Wilson BJ. Comparative Approaches to Studying Strategy: Towards an Evolutionary Account of Primate Decision Making. Evolutionary Psychology. 2013; 11 (3):606–27. PMID: 23864296

70.   Sánchez-Amaro A, Duguid S, Call J, Tomasello M. Chimpanzees, bonobos and children successfully coordinate in conflict situations. Proceedings of the Royal Society of London B: Biological Sciences. 2017; 284:20170259.

71.   de Kwaadsteniet EW, van Dijk E. Social status as a cue for tacit coordination. Journal of Experimental Social Psychology. 2010; 46(3):515–24.

72.   Schelling TC. The Strategy of Conflict. Cambridge (MA): Harvard University Press; 1960.

73.   Grueneisen S, Tomasello M. Children coordinate in a recurrent social dilemma by taking turns and along dominance asymmetries. Developmental Psychology. 2017; 53(2):265–73. https://doi.org/10.1037/dev0000236 PMID: 27819462