Second-and-a-half-party versus third-party punishment

Experimental evidence on punishment behaviour in a public goods game

Jonas Bolzen

September 2018

Abstract

Third-party punishment is a common phenomenon. People are willing to defend the rights of others, even when this is costly for them. However, people might be defending their very own interest even more vigorously. In this paper it is going to be investigated whether or not this is true, by evaluating the results of a public goods game experiment with two treatments. In one treatment, the player with the punishment opportunity had a direct material interest in cooperation. In the other one, the payoff of the punisher was not affected by the cooperation behaviour at all. Due to this design, it is possible to explicitly investigate the respective importance of the material and the emotional dimension for the punishment decision, and the resulting cooperation levels. The results show, that compared to emotions and moral issues, material interests play a minor roll. Third-party punishment is not less common than punishment by actors, who have a direct interest in cooperation.

Content

Tables and figures

1. Introduction

Wherever humans are living in society, they face situations, in which their personal utility collides with the overall utility of a community. A cooperative behaviour would be beneficial for the community, but not cooperating is individually beneficial. These situations appear, where a public good is to be provided or preserved. A good, nobody can be excluded from benefiting from it. The existence of this good requires at the same time cooperation. That means that the individuals must contribute to the good, so that it can be provided. This contribution is costly for them. But for the whole group the benefits of the contribution outweigh the costs, so that cooperative behaviour maximizes the utility of the community.

However, for the individual the benefits from its own contribution do not exceed the costs. On the other hand, the individual can gain utility from the good regardless of its own contribution. Therefore, a self-interested, rational actor is not going to contribute, but gain from the cooperative behaviour of others. This is called free riding. In a group of only self-interested, rational actors, the result is going to be, that nobody contributes and no public good is provided at all. The greater overall utility for the community is not going to be realized.

Situations of such colliding interests are to be found in every society of the world, in small and big scales, in cooperative forms of agriculture and hunting, the preservation of common resources, behaviour in workplace, environmental protection, charity work, and even on markets and in politics. The outstanding relevance of cooperation for human societies urges to a profound investigation of human cooperation behaviour, respectively the influencing factors. An experimental approach to tackle this quest is the Public Goods Game (PGG), an abstract simulation of situations as described above (Camerer und Fehr 2002).

In this experiment, players in a group can decide on their individual amount of contribution. Their individual return of their contribution is always smaller than the costs, but the overall return of the group always larger.

As everyday experience already indicates, not all people behave as self-interested free riders, despite the incentives to do so. Otherwise society would not

exist. And also in PGGs cooperation does occur (Fischbacher et al. 2001; Keser 1996). Several possible motives for this behaviour are discussed in this regard: reciprocity, inequality aversion and certain social norms of fairness and cooperation. In fact, most people seem to be disposed to contribute, provided, that others also contribute. However, the presence of even minor free riding leads to failing cooperation.

Numerous variations of this experiment have been conducted. One approach is to implement punishment opportunities (Camerer und Fehr 2002; Fehr und Gächter 1999). Those enable players of the PGG to punish other players out of their group. However, punishing in such experiments is also costly for the punisher. In many cases people were disposed to bear these costs and vigorously punished those players, who gave less than others, despite the fact, that there were no direct gains from that. Minor contributions are a violation of the motives and values mentioned above, causing negative emotions like the wish for revenge. Furthermore, people might intend to enforce higher contributions, provided the experiment is played over several periods. It turned out, that punishment is indeed able to strongly increase contribution amounts and stabilize them on higher levels (Herrmann et al. 2008; Fehr und Gächter 1999; Fehr und Fischbacher 2004a).

In these designs the punisher has a direct interest in high cooperation, meaning that he gains from contributions. In the real world however, we can observe an additional phenomenon. People often defend the interest of others, even when it is costly for them. They are willing to costly punish uncooperative behaviour, even if they have no direct interest in cooperation in the respective situation. This altruistic punishment is indeed a vital necessity for the functioning of more complex societies, because in these, people which might benefit from cooperation are often not identical with those who are in the position to substantially punish (Fehr und Fischbacher 2003; Fowler 2005). This phenomenon is also investigated by experimental research (Nelissen und Zeelenberg 2009; Charness et al. 2008; Fehr und Fischbacher 2004b). A player just observes the behaviour of other players in games like prisoner dilemma, dictator or investment game. His payoffs are totally independent from this behaviour. He is a *third party*. However, he has the opportunity to punish the other

players. This kind of design is therefore called *third-party punishment* (TPP). In contrast to this, the punishment design described before is called *second-party punishment* (SPP). If he punishes, he loses money and cannot expect to actually gain from that in any way. And furthermore, the motive of revenge is not attributable to him per se. Despite this, the experiments confirm the everyday experience. People are punishing in the TPP design and are thereby able to enforce cooperative behaviour. Several motives are discussed in this regard. Actually, the argumentation goes in a similar direction as in the concern of SPP. The violation of social norms triggers negative emotions (Nelissen und Zeelenberg 2009; Fehr und Fischbacher 2004b). Therefore, people are inclined to avenge those values and altruistically punish.

As mentioned, TPP experiments have been conducted with several game designs, like prisoner dilemma, dictator or investment game. An investigation of TPP in PGGs has – as far as known – not been conducted yet. In this paper, the results of a PGG experiment with TPP are going to be presented.

To create a context, two treatments have been conducted. In the first one a player, who does have interest in cooperation, has the punishment option. You might consider that as a SPP design, but in fact it differs from the previous PGGs with SPP as for example conducted by Herrmann et al. (2008). The treatment is structured in two steps. First an ordinary PGG is going to be played. Following that, an additional player, who did not take part in the actual PGG, has a punishment option. In this regard, it differs from previous experiments. However, despite the fact that he did not participate in the PGG itself, he gains from cooperation. Therefore, it is not TPP either. It lies somewhere between normal SPP and TPP. Therefore, it is going to be called second-and-a-half-party punishment (2,5PP).

Treatment 2 is going to have the TPP design. It is vastly identical with the first one, with the only difference, that the additional player, who has the punishment option, does not gain from cooperation.

Due to this design it is possible to explicitly investigate the importance of material incentives for punishment and the resulting cooperation in a PGG. The comparison of 2,5PP and TPP investigates, whether direct interest in cooperation is a

crucial factor. In contrast, a simple comparison between ordinary SPP and TPP would not be as valid, because various additional factors might distort the results.

2. Experimental design

In this chapter, the design of the experiment of both treatments is going to be described in detail. The two treatments are labelled T1 and T2. Because they are designed largely identical, they are going to be described together and the differences depicted, when they arise. The game was played in groups of four. In each group, there were three players of type A, one of type B. Before the game, every player was randomly assigned to a group and a type. The game went over five periods. Every period was independent concerning endowment and payoff. The group and type assignments however stayed constant during the entire game. There were no practicing periods or a comprehension questionnaire.

2.1. The PGG-step

The first step of every period was an ordinary PGG played by the three players A. Every one of them had an endowment of 10 tokens. They determined individually and simultaneously the number of tokens they wanted to contribute into a common pool and kept the rest. The tokens in the pool were added up and multiplied by a factor of 1.5. After this, the tokens in the pool were equally distributed amongst the three players again. This share plus the number of tokens he retained were the temporary payoff of a player A. The individual contributions and the temporary of every player were then announced to the other players.

2.2. The punishment-step

In the second step of the period, the one player B went into action. He was not taking part in the actual PGG, but the contribution and temporary payoffs were also announced to him. Up to this point, T1 and T2 were structured identically. The difference existed in the endowment and the further gaining options of B.

2.2.1. Treatment 1

In T1 B started with an initial endowment of 15. Furthermore, he received the same number of tokens every player A received out of the common pool. These two amounts added up to his temporary payoff, which therefore depends on the behaviour of the three players A, which makes T1 the 2,5PP treatment.

2.2.2. Treatment 2

In T2 B received just a set endowment and could not gain any further tokens. Therefore, he did not have a temporary payoff depending on the behaviour of the players A, but merely his endowment.

It would not have been legit to simply determine a specific constant endowment for all players B in T2, for example 15. In this case the endowment of B in T2 would have systematically differed from the amount of tokens that were available for B in the T1, the temporary payoff. To prevent this, the endowments of every specific B in T2 reflected the temporary payoff of one specific B in T1. Therefore, every B in T2 was paired with a B in T1. The temporary payoffs of the B in T1 were implemented as the initial endowments of the B in T2. This was done period specific. Therefore, the initial endowments of B in T1 can be considered as statically equal to the temporary payoffs of B in T2. The procedure is exemplary illustrated in table 1.

Therefore, the difference between T1 and T2 was, that in T1 B received a fixed endowment and might have gained further tokens out of the pool depending on the behaviour of the players A, while in T2 B received an alternating but set endowment that was not depending on the behaviour of the players A and no further gains were possible. The players in T2 were not told how this alternation arose. They were only told that they would receive changing endowments and that those are not depending on the behaviour of the players A. Therefore, T2 is the TPP treatment.

2.2.3. Punishment and final payoff

The rest of the game was again equal for both treatments. B could now selectively assign punishment points (PP) to the players A. Every PP cost B one token, but the A who received the PP lost three tokens. B could assign up to five PP to each A. After the assignment of the PP, the final payoffs were announced. The final payoffs of the players A were their temporary payoffs minus the received PP times three. The final payoff of B was his temporary payoff, respectively his initial endowment minus the assigned PP. Those were the final payoffs of one period. Negative payoffs were not possible and considered as being zero. Not integer values that emerged in the game had been rounded.

Table 1

Period		**1**	**2**	**3**	**4**	**5**
T1	**Initial endowment**	**15**	**15**	**15**	**15**	**15**
	Tokens received out of the pool	**5**	**7**	**7**	**10**	**8**
	Temporary payoff	**20**	**22**	**22**	**25**	**23**
T2	**Initial endowment**	**20**	**22**	**22**	**25**	**23**

Note: Exemplary illustration of the design of endowments for a player B

3. Theoretical solution

This chapter expounds the game's theoretical solution of the experiment as described in the chapter above for both treatments.

3.1. The PGG-step

As mentioned in chapter 1, the dominant strategy in a PGG is to contribute nothing. The individual return of a token contributed by a player A arises out of the multiplication by the factor 1.5 and the dividing between the three players A. The result of 0.5 is lower than one. The overall payoff of the group would be however maximized, when every player A contributes his whole endowment, because the group's return of a contributed token is larger than one. It is 1.5 for T2 and 2.0 for T1, because in T1 B receives the same amount as every player A received out of the pool.

3.2. The punishment-step

3.2.1. Treatment group

The endowment and therefore the payoff of B is affected by the behaviour of the players A. He has an interest in contributions. Nevertheless, towards backward induction it can be shown that punishing is an implausible threat. Therefore, B is not punishing, and A is not changing his behaviour – meaning, he still contributes nothing – due to the existence of the punishment option.

3.2.2. Control group

B has no interest in the behaviour of the players A, because it does not affect his payoff. Therefore, he is not bearing the cost of punishing them. Anticipating that, A is not changing his behaviour. All in all, according to the theoretical solution, in T1 as well as in T2 no player A contributes, and B does not punish.

4. Hypotheses

This chapter is dedicated to the derivation of hypotheses regarding the results of the actual experiment. First of all, it is highly expected that punishment and contribution are taking place at all in both treatments, despite the described theoretical solution. These minor hypotheses are based on various prior experiments depicted in chapter 1.

However, most interesting is, of course, in which way the results differ between the treatments. As pictured above, the only crucial variation between the treatments was the fact, that the payoff of B was directly affected by the behaviour of the players A in the T1, but not in T2. On that basis, it is going to be elaborated, how this difference might have influenced the results. Two main differences might arise: a different punishment behaviour of B and a different contribution behaviour of A.

4.1. Hypothesis concerning the punishment behaviour of B

As mentioned in chapter 1, several motives might be important for a B. Specifically a B in T1 (2,5PP) might face the following motives: First, the more rational one of aiming to increase his own payoff. By punishing he might try to enforce higher contributions and therefore a higher payoff for himself. As described, punishment is indeed able to realize higher contribution levels.

Furthermore, moral, respectively emotional motives might play a role, like social norms of fairness and cooperation, inequality aversion, positive reciprocity and revenge. Revenge could obviously be important, when B wants to punish an A who gave very few tokens in the pool. But positive reciprocity might also emerge when it is coming to punishment. B might consider relatively high contributions of an A as a nice act and therefore behaves reciprocal to him by avenging him and punishing an A, who was enriching himself at the expense of the *nice* A by contributing less.

These motives might induce a B in T1 to punish an A. In T2 (TPP), however, B does not face all of those motives. Because of the fact, that he does not materially gain from contributions, the motives of increasing his own payoff, positive

reciprocity and revenge per se are no longer attributable to him. The motives left are inequality aversion and social norms of fairness and cooperation. As described in chapter1, those motives can trigger strong negative emotions and thereby lead to altruistic punishment. However, regardless of the respective actual importance of each motive for the final behaviour, it can be assumed, that a B in T2 – facing less possible motives – is punishing less than a B in T1 – facing additional possible motives. This is the first main hypothesis.

This different punishment behaviour might arise in two dimensions. First, the absolute amount of punishment is higher in T1 and second, punishing in T1 is implemented relatively more strictly. The second one is based on the premise, that a player A is punished higher the lower his contribution is. This increase of punishment for lower contributors is expected to be higher in T1 than in T2.

4.2. Hypothesis concerning the contribution behaviour of A

The second main hypothesis is largely based on the first one. If the first one turns out to be true, it can be assumed, that the more vigorous punishment increases the contribution levels, like it was observed in prior experiments.

Furthermore, similar to the reflections in 4.1, an A might already anticipate a harsher punishment behaviour by B in T1, as he would in T2. Therefore, he might have more fear of punishment for minor contributions. The threat appears to be more plausible. An A in T1 might further anticipate, that B is not running out of steam as quickly as an A in T2 might think. B in T1 is less expected by A to give up punishing quickly, because he is fighting for his very own. These imaginations of A about the motives and behaviour of B can be assumed to increase contributions.

Furthermore, this hypothesis is reaffirmed by the fact, that the overall return of one contributed token for the group is larger in T1 (2.0) than in T2 CG (1.5) (see chapter 2). The higher benefit of contributions for the group might be a motive for A to increase them. This last motive argues in a more philanthropic direction, than those elaborated before.

To sum up, the second main hypothesis is, that A is contributing more in T1. The two dimensions of this hypothesis are: First, the total contribution levels are

higher. Second, due to the expected higher punishments and their contribution encouraging effect, the development of the contribution levels over time are differing, meaning contribution levels are experiencing a larger increase over the periods in T1 compared to T2. Both main hypotheses are highlighted in table 2.

Table 2, Hypotheses

Hypothesis 1: Higher punishment in T1 (2,5PP) than in T2 (TPP)
1.1. Higher absolute punishment
1.2. Higher punishment relative to contribution

Hypothesis 2: Higher contribution in T1 than in T2
2.1. Higher absolute contribution levels
2.2. Higher increase over time

5. Empirical results

In this chapter – after a brief presentation of the conditions, in which the experiment was conducted, and the pool of participants – the empirical results are going to be elaborated and the validity of the two main hypotheses is going to be checked.

5.1. Conditions and participants of the experiment

The experiment was conducted in the context of a master seminar at the University of Passau in the timespan of the 18th to 26th June 2018 in a computer lab of the university. The participants were recruited on the campus verbally and by posters. The recruitment was supported by the promise of having coffee and snacks for every participant.

Ten sessions of each treatment took place with four to twenty participants per session. Each session took approximately a quarter of an hour. The experiment was conducted on computers using the software z-Tree by Fischbacher (2007). The participants played each in a separated booth. Every actions and data of the participants were registered anonymously.

There were 264 participants in 66 groups. 132 participants in 33 groups took part in each treatment. Each participant took only part in one session. Most of them were students of the university with very few exceptions, enrolled in over 10 different courses of study. Further pool information is given in table 3.

The players were informed, that due to the explicitly didactic nature of the experiment, there would be no payoff in real money possible. But they were asked to imagine, that they would play for real money and try to behave according to that.

Table 3, Pool Data

	In total	T1	T2
Participants	264	132	132
Groups	66	33	33
Female	168 (63,64 %)	80 (60,6%)	88 (66,67%)
Average age	22,6	22,72	22,48

5.2. First hypothesis: Punishment behaviour of B

The first hypothesis is, that B is punishing more in T1, than in T2. Before checking for that, it should be mentioned, that – deviant from the theoretical solution – there was punishing in both treatments. This confirms results of prior experiments.

The first dimension of the first hypothesis is, that the total amount of punishment is higher in T1 than in T2. The total number of PP assigned in one period to all of the three players A was 3.461 on average in T1, which is slightly less than in T2 with an average of 3.636. Furthermore, the development of the punishment behaviour over time was very similar in both groups, meaning that it stayed mostly unchanged (see figure 1).

Table 4 displays the results of a corresponding OLS approach. It shows, that the players B in T1 and T2 did not significantly differ in this regard (column 1). This holds true for taking the periods in account (column 2 & 3). Due to these results, the first dimension must be rejected.

Furthermore, table 5 displays the connection between PP assigned to specific players and their contribution behaviour. In contrast to the previous approach, the dependant variable is now the amount of PP a specific player received in a period, and not how much PP are assigned by B in total in a period. In both groups those players who did contribute less received more PP in the following punishment step. These results confirm the premise of the second dimension of the first hypothesis. However, the extent of this correlation hardly differs between the groups. The second dimension does not hold true, too.

According to that, the first hypothesis must be rejected. B did not punish more in T1 than in T2. The punishment behaviour is approximately equal.

Figure 1

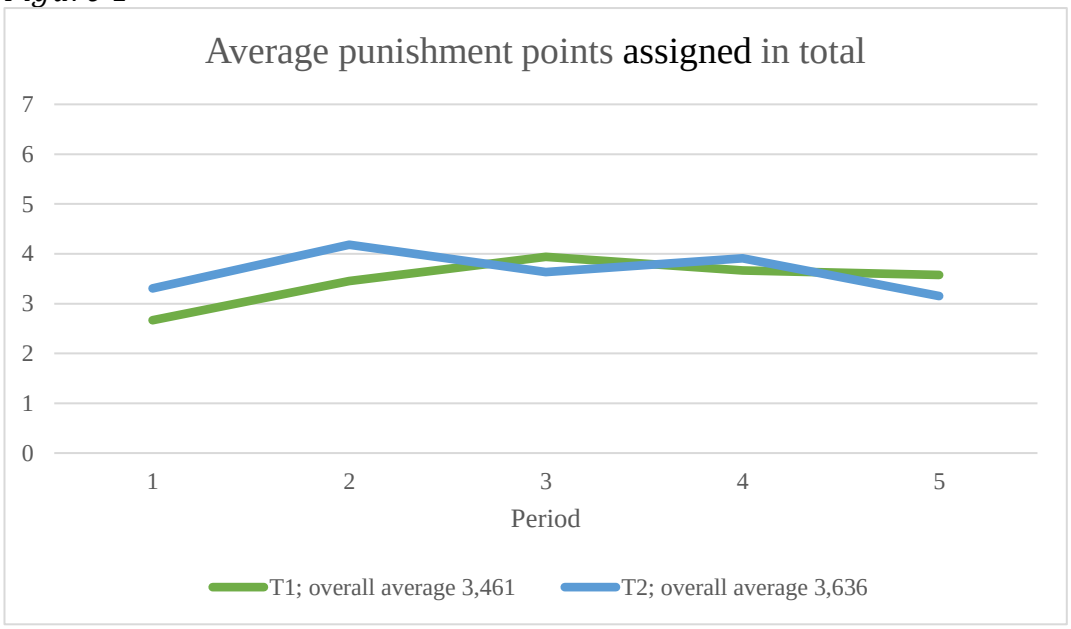

Average punishment points assigned in total
7
6
5
4
3
2
1
0
1
2
3
4
5
Period
T1; overall average 3,461
T2; overall average 3,636

Table 4, Punishment behaviour 1

	Dependent variable:			
	Punishment points assigned in total			
	(T1&T2)	(T1)	(T2)	(T1&T2)
Treatment	-0.1758			-0.9576
	(0.3783)			(0.8883)
Period		0.2030	-0.0576	-0.0576
		(0.2049)	(0.1725)	(0.1894)
Treatment x Period				0.2606
				(0.2678)
Constant	3.6364***	2.8515***	3.8091***	3.8091***
	(0.2675)	(0.6794)	(0.5722)	(0.6281)
Observations	330	165	165	330
R^2	0.0007	0.0060	0.0007	0.0044
Adjusted R^2	-0.0024	-0.0001	-0.0054	-0.0047
Residual Std. Error	3.4363 (df = 328)	3.7214 (df = 163)	3.1341 (df = 163)	3.4403 (df = 326)
F Statistic	0.2158 (df = 1; 328)	0.9822 (df = 1; 163)	0.1114 (df = 1; 163)	0.4857 (df = 3; 326)

*Note: *p<0.1; **p<0.05; ***p<0.01. The dependent variable is the number of punishment points assigned in total by a specific B in a specific period. Column 1 and 4 depict this variable for both treatments, column 2 for T1 and column 3 for T2. The explanatory variables are the treatment as a dummy variable, the period and the interaction term between them.*

Table 5, Punishment behaviour 2

	Dependent variable:		
	Punishment points assigned to a specific A		
	(T1)	(T2)	(T1&T2)
Treatment			-0.3700*
			(0.1971)
Contribution	-0.1314***	-0.1814***	-0.1814***
	(0.0217)	(0.0181)	(0.0196)
Treatment x Contribution			0.0500*
			(0.0282)
Constant	1.9682***	2.3382***	2.3382***
	(0.1510)	(0.1271)	(0.1377)
Observations	495	495	990
R^2	0.0693	0.1689	0.1149
Adjusted R^2	0.0674	0.1672	0.1122
Residual Std. Error	1.5290 (df = 493)	1.3182 (df = 493)	1.4275 (df = 986)
F Statistic	36.7144*** (df = 1; 493)	100.2030*** (df = 1; 493)	42.6614*** (df = 3; 986)

*Note: *p<0.1; **p<0.05; ***p<0.01. The dependent variable is the number of punishment points assigned to a specific A in a specific period. Column 1 depicts this variable for T1, column 2 for T2 and column 3 for both treatments. The explanatory variables are the contributions of a specific A in the same period and interaction term between treatment and contribution.*

5.3. Second hypothesis: Contribution behaviour of A

The second hypothesis is, that A is contributing more in T1 than in T2. Also in regard to contributions, players did not behave as the theoretical solution demands. There was contribution in both treatments, as there was in prior experiments.

The first dimension of the second hypothesis is, that the total contribution levels are higher in T1 than in T2. A was contributing 6.2 on average in T1, and nearly the same amount of 6.208 in T2. The second dimension is, that contribution levels are increasing more in T1 than they do in T2. However, the development over time was very similar, meaning that it was largely constant in both treatments (see figure 2).

An OLS regression (table 6) shows, that there are no statistical differences between the groups regarding the aggregated contribution behaviour (column 1) and the development over time (column 2, 3 & 4). Both dimensions and therefore the second hypothesis must be rejected. Contribution behaviour in T1 is not higher than in T2, but largely the same.

Figure 2

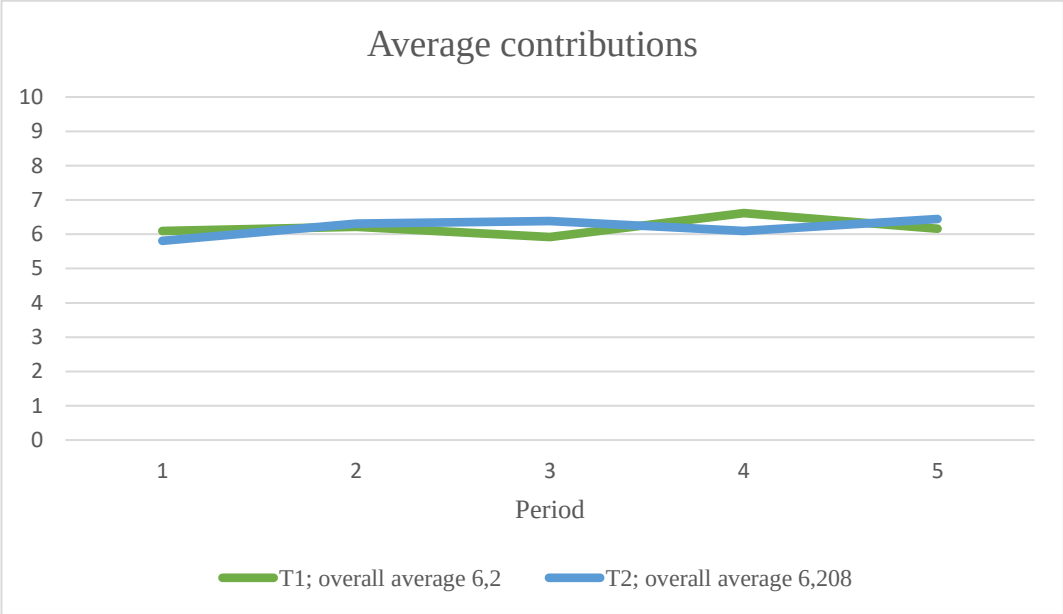


Table 6, Contribution behaviour

	Dependent variable:			
	Contribution			
	(T1&T2)	(T1)	(T2)	(T1&T2)
Treatment	-0.0081			0.1434
	(0.2049)			(0.4806)
Period		0.0545	0.1051	0.1051
		(0.1009)	(0.1040)	(0.1025)
Treatment x Period				-0.0505
				(0.1449)
Constant	6.2081***	6.0364***	5.8929***	5.8929***
	(0.1449)	(0.3346)	(0.3450)	(0.3398)
Observations	990	495	495	990
R^2	0.000002	0.0006	0.0021	0.0014
Adjusted R^2	-0.0010	-0.0014	0.00004	-0.0017
Residual Std. Error	3.2230 (df = 988)	3.1745 (df = 493)	3.2728 (df = 493)	3.2240 (df = 986)
F Statistic	0.0016 (df = 1; 988)	0.2923 (df = 1; 493)	1.0200 (df = 1; 493)	0.4453 (df = 3; 986)

*Note: *p<0.1; **p<0.05; ***p<0.01. The dependent variable is the number of contributed tokens by a specific A in a specific period. Column 1 and 4 depict this variable for both treatment, column 2 for T1 and column 3 for T2. The explanatory variables are the treatment as a dummy variable, the period and interaction term between both.*

6. Discussion

The purpose of the experiment was to investigate the importance of a material interest in contributions by for punishment and the resulting contributions in a cooperation game, the PGG. The initial hypotheses were, that because of the absence of such an interest, that punishment and cooperation would be less in the TPP treatment. It turned out, that this is not the case. In fact, the results do not depict any substantial differences between the treatments regarding punishment or contribution behaviour.

As mentioned, various motives must be taken in account for the punishment decision. Furthermore, it was argued, that a B in T1 would face more motives. If that is true, the results lead to the conclusion, that the material matters are much less important than the moral or emotional ones. Therefore – with some restrictions discussed below – it was shown, that the negative emotions triggered by uncooperative behaviour and the accompanying violation of certain values seem to be the crucial motivation, whereas material issues play a minor roll.

An alternative, respectively supplementary explanation is, that due to the lack of material incentives in T2, an additional motive arose, the warm glow of altruism. In T2, player B was able to consider his punishing as a pure act of justice and knew, that the others would consider it like this, too. He was able to think that his behaviour was not caused by “low” motives like money, but only by altruism. This untarnished motive is only possible in T2 and may have influenced his behaviour.

The hypothesis, that the contribution levels in T1 would be higher than in T2 was largely based on the first hypothesis regarding punishment. Therefore, the fact, that contribution did not differ between treatments does not surprise. On the other hand, it was argued, that already the imaginations of A above B´s behaviour – meaning that A fears punishment more in T1 – may increase contributions. That was not the case. It would be very enriching to investigate further in this regard. Fehr und Fischbacher (2004b) did so by specifically asking for the respective punishment expectations in a TPP game, but without comparison to a different punishment design.

Also, it did not seem to matter for A, that the group's overall return of his contribution was larger in T1 than in T2 (2,0 versus 1,5). Despite the fact, that this difference was rather subtle, a possible explanation could be, that an A might not consider B as really deserving this money, because he was not participating in the PGG and did not contribute anyway. His gains from the PGG therefore might be considered as of minor importance for A.

Furthermore, certain restrictions of the experiment must be taken in account. First, it is quite complicated. It was played without practicing rounds. Therefore, people might have needed some time to really understand the indications of the design. That is not a problem per se, but due to the circumstances of the experiment, there were only five periods. Most prior experiments had at least ten periods (Herrmann et al. 2008; Fischbacher et al. 2001; Fehr und Gächter 1999). Especially considering the development of contributions over time, it would have been favourable to conduct more periods.

A second problem is the size of the sample. 264 participants are quite substantial. However, only a quarter of them were players of type B, meaning, that only 33 players B participated in each treatment. Because of the importance of the punishment behaviour in this design, this is not that much.

Third, there were no actual payoffs. Of course, that holds true for both treatments. However, the only difference between the treatments were the – hypothetic – material incentives for punishers in T1. In contrast to the still valid moral and emotional motives, including certain restrictions, the material factor might become too subtle for strong results in this design, especially considering the small sample.

7. Conclusion

The experiment described and evaluated in this paper allows to explicitly depict the relevance of a material interest in cooperation by the punisher in a PGG. One treatment was conducted with implementing such an interest, one without. The results did not show any differences regarding punishment or contribution behaviour. Therefore, it is concluded, that the material motives are of minor importance compared to the emotional and moral inducements. Third party punishment seems to be a phenomenon that is similarly common and powerful as punishment by an actor who has an interest in cooperation.

Literature

Camerer, Colin F.; Fehr, Ernst (2002): Measuring social norms and preferences using experimental games. A guide for social scientists. Working Paper No. 97. In: *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*, S. 55–95.

Charness, Gary; Cobo-Reyes, Ramón; Jiménez, Natalia (2008): An investment game with third-party intervention. In: *Journal of Economic Behavior & Organization* 68 (1), S. 18–28.

Fehr, Ernst; Fischbacher, Urs (2003): The nature of human altruism. In: *Nature* 425 (6960), S. 785.

Fehr, Ernst; Fischbacher, Urs (2004a): Social norms and human cooperation. In: *Trends in cognitive sciences* 8 (4), S. 185–190.

Fehr, Ernst; Fischbacher, Urs (2004b): Third-party punishment and social norms. In: *Evolution and human behavior* 25 (2), S. 63–87.

Fehr, Ernst; Gächter, Simon (1999): Cooperation and punishment in public goods experiments.

Fischbacher, Urs (2007): z-Tree. Zurich toolbox for ready-made economic experiments. In: *Experimental economics* 10 (2), S. 171–178.

Fischbacher, Urs; Gächter, Simon; Fehr, Ernst (2001): Are people conditionally cooperative? Evidence from a public goods experiment. In: *Economics Letters* 71 (3), S. 397–404.

Fowler, James H. (2005): Altruistic punishment and the origin of cooperation. In: *Proceedings of the National Academy of Sciences* 102 (19), S. 7047–7049.

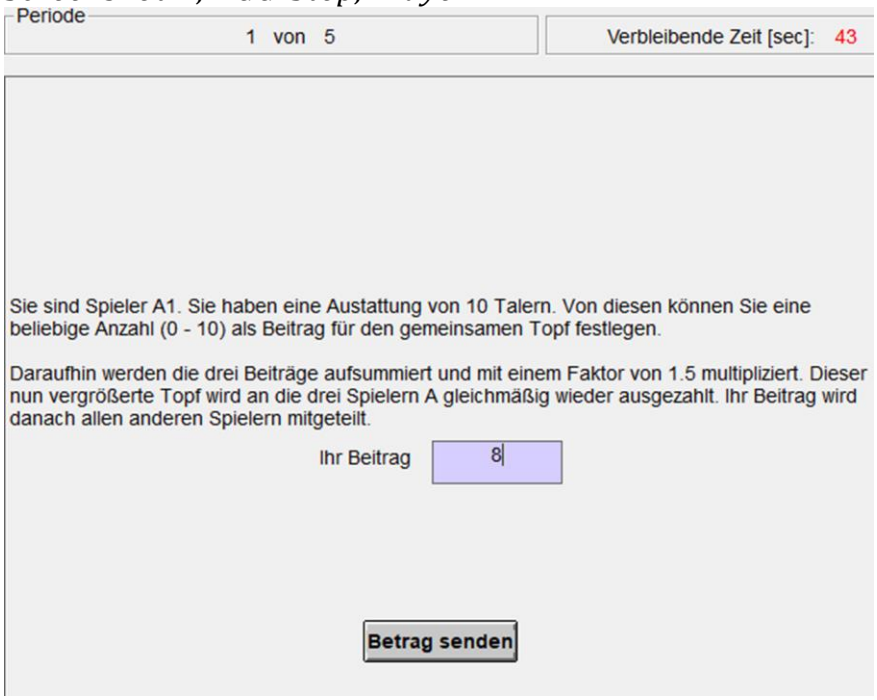
Herrmann, Benedikt; Thöni, Christian; Gächter, Simon (2008): Antisocial punishment across societies. In: *Science* 319 (5868), S. 1362–1367.

Keser, Claudia (1996): Voluntary contributions to a public good when partial contribution is a dominant strategy. In: *Economics Letters* 50 (3), S. 359–366.

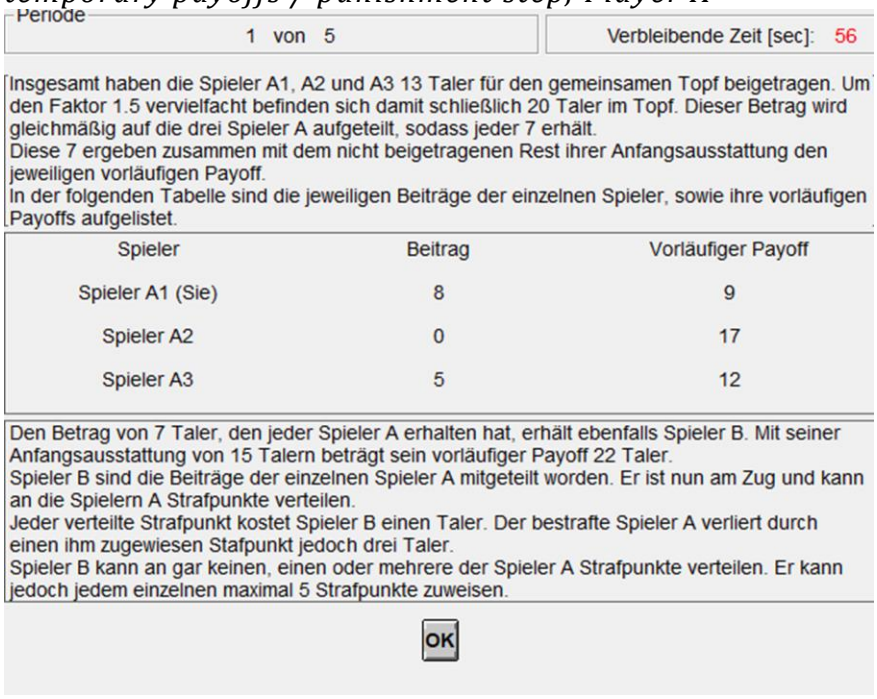
Nelissen, Rob M. A.; Zeelenberg, Marcel (2009): Moral emotions as determinants of third-party punishment. Anger, guilt and the functions of altruistic sanctions. In: *Judgment and Decision making* 4 (7), S. 543.

Appendix

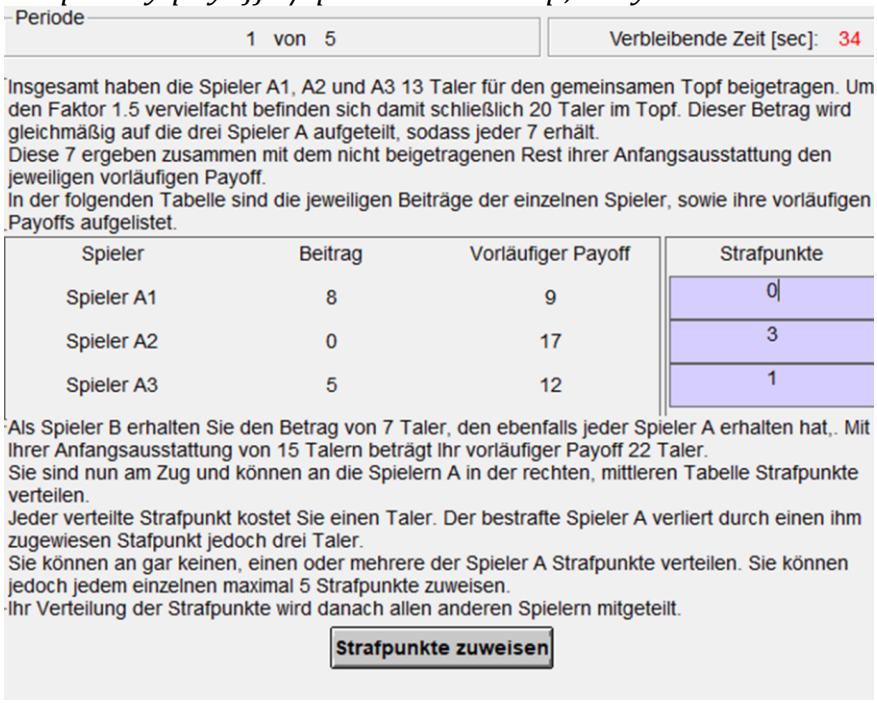
Screenshot 1; PGG-Step, Player A



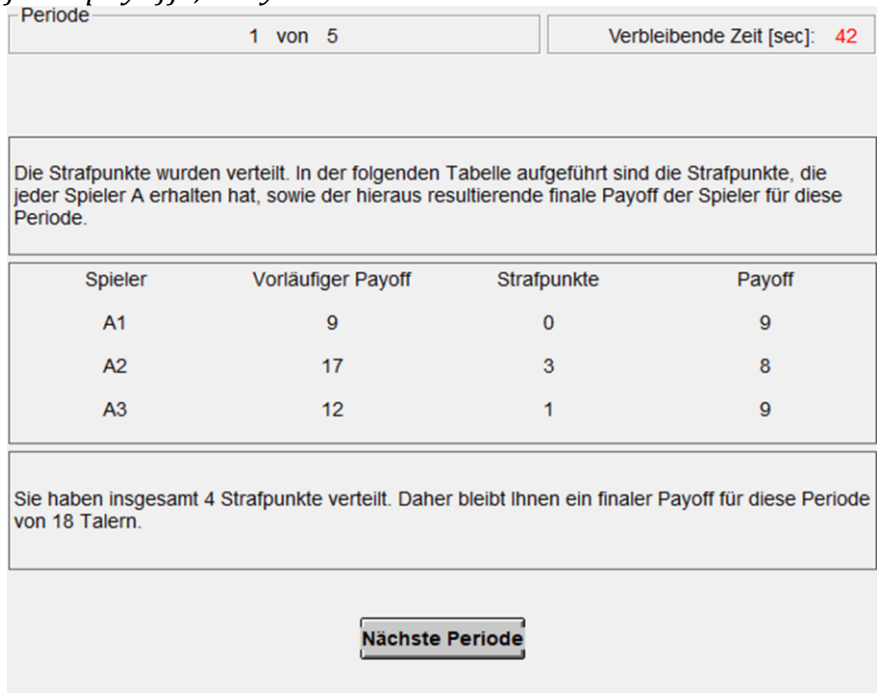
Screenshot 2; Announcement of contributions and temporary payoffs / punishment step, Player A



Screenshot 3; Announcement of contributions and temporary payoffs / punishment step, Player B



Screenshot 4; Announcement of punishment points and final payoffs, Player B



Game instructions for both treatments: The highlighted paragraphs are respectively referring to one the treatments; Green -> T1, Blue -> T2

Sie spielen mit drei anderen Spielern in einer Gruppe. Es werden zwei Rollen unterschieden:
-Drei Spieler mit der identischen Rolle A. Zu Unterscheidung werden sie als A1, A2 und A3 bezeichnet.
- Ein Spieler B.

Vor Beginn des Spiels werden die Teilnehmer in Gruppen eingeteilt und es wird jedem eine Rolle zufällig zugewiesen.
Das Spiel läuft über 5 Runden (Perioden). Jede der Perioden ist bezüglich Ausstattungen und Payoffs unabhängig von den vorhergehenden. Die Gruppeneinteilung und Rollenzuweisung bleibt jedoch unverändert.

Spieler A

Die drei Spieler A sind zuerst am Zug. Jeder von ihnen hat eine Anfangsausstattung von je 10 Talern.Ihre Taler können die Spieler beliebig in einen gemeinsamen Topf investieren. Das bedeutet sie können entweder gar nichts, einen Teil ihrer Ausstattung oder alles in diesen Topf investieren. Dies stellt ihren jeweiligen Beitrag dar. Die Entscheidung über die Beitragshöhe treffen die Spieler zeitgleich voneinander.
Daraufhin werden die drei Beiträge aufsummiert und mit einem Faktor von 1,5 multipliziert. Dieser nun vergrößerte Topf wird an die drei Spielern A gleichmäßig wieder ausgezahlt. Jeder erhält ein Drittel, unabhängig davon, wie hoch sein Beitrag war.
Die Auszahlung aus dem Topf für einen Spieler, plus der Rest von seiner Anfangsausstattung ergibt seinen vorläufigen Payoff.
Die jeweiligen Beiträge der Spieler, sowie ihr vorläufiger Payoff werden daraufhin allen anderen Spielern mitgeteilt.

Spieler B

Jetzt ist der eine Spieler B am Zug. **Er verfügt über eine Anfangsausstattung von 15 Talern. Zusätzlich erhält er nochmal genauso viele Taler, wie auch die Spieler A jeweils aus dem gemeinsamen Topf bezogen haben. Beides zusammen ist der vorläufige Payoff des Spielers B.**

Seine Ausstattung wechselt von Periode zu Periode. Mal ist sie höher, mal niedriger. Sie ist jedoch völlig unabhängig vom Verhalten der anderen Spieler.

Dem Spieler B werden ebenfalls die Beiträge der Spieler A mitgeteilt. Er kann daraufhin Strafpunkte an die einzelnen Spieler A verteilen.

Ein verteilter Strafpunkt kostet Spieler B einen Taler. Der Spieler A, der bestraft wird, verliert durch jeden Strafpunkt jedoch drei Taler.

Spieler B kann an gar keinen, einen oder mehrere der Spieler Strafpunkte verteilen. Er kann jedoch jedem einzelnen maximal 5 Strafpunkte zuweisen.

Payoff

Der schlussendliche Payoff der Periode für Spielers B ergibt sich aus seinem vorläufigen Payoff, minus den Kosten für die verteilten Strafpunkte.

Der schlussendliche Payoff der Periode für Spieler B ergibt sich aus seiner Ausstattung, minus den Kosten für die verteilten Strafpunkte.

Der schlussendliche Payoff der Periode für einen Spieler A ergibt sich aus seinem vorläufigen Payoff, minus den Abzügen durch erhaltene Strafpunkte.

Die verteilten Strafpunkte und die Payoffs werden nun allen Spielern mitgeteilt.

Negative Payoffs sind nicht möglich. Diese werden als Null gewertet. Wenn sich im Spiel nichtganzzahlige Werte bilden, werden diese nach den üblichen Regeln gerundet. Damit ist eine Periode beendet und die nächste startet.