HOW PEER-PUNISHMENT AFFECTS COOPERATIVENESS IN HOMOGENEOUS AND HETEROGENEOUS GROUPS

Marianna Bicskei, Matthias Lankau, Kilian Bizer

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

HOW PEER-PUNISHMENT AFFECTS COOPERATIVENESS IN HOMOGENEOUS AND HETEROGENEOUS GROUPS

A Public Goods Experiment with Social Identity

Marianna Bicskei, Matthias Lankau* and Kilian Bizer

Economic Policy and SME Research, University of Göttingen,

Platz der Göttinger Sieben 3, 37073 Göttingen, Germany

Abstract

This article analyzes how the anticipation of peer-punishment affects cooperativeness in the provision of public goods under social identity. For this purpose we conduct one-shot public good games with induced social identity and implement in-group, out-group and random matching protocols. Our measure of cooperativeness is subjects' conditional contribution elicited via the strategy method, which allows for observing behavior contingent on every possible level of group members' cooperation. We demonstrate, firstly, that the social environment is a determinant of how the threat of peer-punishment influences cooperation. The strongest increase is clearly evident when subjects interact with members of different identities, which is especially the case for individuals who were initially categorized as free-riders. Secondly, anticipation of peer-punishment clearly eliminates the typically existing in-group bias without punishment and renders out-group members to be as cooperative as in-groups members. Lastly, the results indicate that the institutions of peer-punishment and social identity may be complemented in order to raise subjects' cooperativeness.

(*JEL* C92, D03, D73, H41)

* Corresponding Author. Tel: +4955139174890, E-mail addresses: matthias.lankau@wiwi.uni-goettingen.de (M. Lankau), marianna.bicskei@wiwi.uni-goettingen.de (M. Bicskei), bizer@wiwi.uni-goettingen.de (K.Bizer).

I.   INTRODUCTION

When a public good is provided, the social optimum is reached if all group members bear the costs and contribute to its provision. However, since no one can be excluded from the benefits the public good generates there is a strong incentive for self-interested individuals to free-ride on their peers' contributions. Although free-riding does not reveal itself as prominent a behavioral pattern as classic economic theory predicts, experimental evidence suggests that there is still a considerable lack in cooperativeness lowering social efficiency. Nevertheless, previous research showed that individuals have a proclivity to adjust their behavior depending on the institutional environment in which the public good is provided. In this respect, two factors are of crucial importance. Firstly, social identity of group members with whom individuals interact influences cooperation. Recently, Lankau et al. (2012) demonstrated that individuals reveal the preference for consistently higher conditional cooperation when matched with group members with a common identity (in-group matching) as opposed to being matched with individuals saliently belonging to different social groups (out-group matching) or randomly matched individuals. This in-group bias in the propensity to cooperate seems to be a deeply rooted phenomenon of human interactions (Eaton, Eswaran, and Oxoby 2011) and has been confirmed by numerous contributions both in psychological and economic experimental research (cf. for instance Taifel and Turner 1979; Chen and Li 2009). Secondly, it is a well-established phenomenon that the possibility of costly decentralized punishment by individual group members for uncooperative behavior strongly enhances cooperation (cf. Fehr and Gächter 2000; Masclet et al. 2003). Clearly, both institutions impact on people's cooperativeness and are thus highly relevant to the maintenance of common resources. However, their interaction in a public good environment did not receive any academic

attention, so far. Consequently, the goal of this article is to provide initial evidence on subjects´ cooperativeness in the provision of public goods in case their social identity is salient and there is a possibility to punish group members for misbehavior. In particular, we ask how the institution of costly peer-punishment impacts subjects' cooperativeness in homogeneous groups composed of members sharing a common social identity and in heterogeneous groups consisting of members of different identities. For this purpose, we artificially induce social identity with the help of a simple group task that subjects solve jointly by communicating anonymously via chat. Our experiment is based on one-shot public good games using the strategy method (Fischbacher, Gächter, and Fehr 2001; Selten 1967), which allows us to observe contributions to the public good conditional on every possible average contribution level of the remaining group members. It thus elicits subjects' cooperativeness in more detail than unconditional contributions alone, and serves us as a preferred measure of cooperativeness. What is more, it enables us to classify subjects into different cooperation types depending on their initial cooperation strategy providing deeper insights into cooperative adjustments under different institutional settings.

In summary, our findings provide hints that group composition in terms of individuals´ identity decisively influences to what extent subjects alter their cooperativeness under punishment compared to equal matching conditions that are free of peer-sanctions. In particular, we prove that subjects who are matched with out-group members increase their cooperativeness to the highest degree. Based on the example of free-riders we are able to show that this is predominantly caused by an anticipation of comparatively strong punishment for uncooperative behavior by group members of different identities. What is more, under punishment subjects' cooperativeness is equal within in- and out-groups essentially revealing that peer-punishment eliminates the in-group bias that is present without peer-punishment.

The remainder of the paper is structured as follows: Section 2 deals with the relevant literature and presents our hypotheses. While Section 3 introduces the experimental design, Section 4 discusses the relevant findings. The article ends with a conclusion in Section 5.

II.   LITERATURE REVIEW AND RESEARCH HYPOTHESES

When people face social dilemmas, individual interest is at odds with social interest. Thus, much theoretical and empirical research has been devoted to understand how to enhance voluntary contributions to public goods. Research in experimental economics demonstrated that social identity as "the individual's self-concept derived from perceived membership in social groups" (Charness, Rigotti, and Rustichini 2007, 1342) and also peer-punishment are prone to foster cooperativeness in the provision of public goods.

To begin with, the experimental economic research on social identity is rooted in the social identity theory introduced by Tajfel and Turner (1979). The core element of this theory is the analysis of the psychological foundations of discriminatory behavior, which is determined through three processes. Accordingly, individuals relatively quickly sort themselves by certain social categories such as gender or race (categorization) and derive self-esteem from that (in-)group (identification). These processes are subsequently complemented by individuals comparing their in-group with out-groups they do not identify with (comparison). Taken together, these processes generally trigger in-group favoritism and thus out-group discrimination commonly referred to as *in-group bias* (cf. Tajfel and Turner 1986; Hoff and Pandey 2006).[1] By now, there are several studies that focus on the effects of social identity on

---

[1] See Chen and Li (2009) for a detailed account on theoretical and empirical findings on social identity theory.

subjects' cooperativeness in a public good context. Firstly, Eckel and Grossmann (2005) study the impact of team identification on cooperation in a repeated-play public good game with various degrees of enhanced team identification. The authors find that strengthening team identification (e.g., prior group task before the game, creation of in-group/out-group conflict) yields consistently higher cooperation levels and less free-riding. Secondly, Lankau et al. (2012) explicitly study the impact of social identity on cooperation preferences in public goods provision. Using the strategy method in multiple one-shot public good games, the authors confirm the existence of an in-group bias. Specifically, they find that when subjects interact with members of their own identity they show a consistently higher preference for conditional cooperation and thus less self-serving bias[2] than when matched with out-group and random individuals. What is more, even the least socially oriented cooperation types, initially identified as free-riders, reveal higher levels of conditional cooperation in in-group matching. Other recent studies documenting subjects' increased cooperativeness in providing public goods when bound together by a common identity include Blackwell and McKee (2013), Chakravarty and Fonseca (2012 and 2013). Similarly, various other experiments outside the public goods context demonstrate that social identity positively influences social welfare when individuals interact with in-group members as opposed to out-group members (cf. Chen and Li 2009; Charness, Rigotti, and Rustichini 2007; McLeish and Oxoby 2011; Goette, Huffman, and Meier 2006; Chen and Chen 2011).[3] In sum, the existing experimental research provides solid evidence that salient social identity in terms of group composition

---

[2] Subjects reveal self-serving bias when they do not reciprocate their team members' contributions by an equal own contribution (perfect conditional cooperation), and depart in the selfish direction.
[3] For a detailed review of these studies, see Lankau et al. (2012).

matters with regard to cooperativeness, and confirms the well-established in-group bias in human interactions.[4]

The possibility of costly decentralized punishment by individual group members without central authority has been identified as an effective instrument to maintain cooperation in the provision of public goods, as well. Initially, Fehr and Gächter (2000) studied the punishment institution in an environment where both the offender and the punisher bear tangible costs of the penalties, though one punishment point costs the offender three times more than the punisher. Specifically, the authors conducted two treatments with and without punishment and distinguished between stranger[5] and partner matching.[6] In both settings without punishment constraints they found that contributions to the public good decreased over the ten periods approaching almost complete free-riding. On the contrary, contribution levels steadily increased towards full cooperation under punishment in the partner matching condition.[7] In another experiment Fehr and Gächter (2002) have identified similar tendencies for strangers, as well.[8] Contribution levels increased under punishment, while they decreased when the opportunity to punish was removed. Likewise, Masclet et al. (2003) and Noussair and Tucker (2005) demonstrated that the fear of being punished drives subjects to increase their contribution to the public good. What is more, several other studies have replicated the experiments of Fehr and Gächter and by and large confirmed their findings (cf. Anderson and

---

[4] For further research on in-group favoritism see Hermann et al. (2012).
[5] Stranger matching meant a random group assignment in each period.
[6] The composition of the group was stable over the periods.
[7] Full cooperation was declared as a dominant behavioral standard regardless whether the punishment or the non-punishment treatment was conducted firstly. However, this behavioral norm was not found in the stranger treatment.
[8] This experiment was designed identically to the previous experiment (Fehr and Gächter 2000) with the exception that punishment and non-punishment treatments were conducted only with strangers. In this study a subject could never meet another subject more than once during the game (perfect stranger matching).

Putterman 2006; Bochet et al. 2006; Nikiforakis 2008; Page et al. 2005; Sefton et al. 2007).

All in all, these findings confirm that costly peer-punishment motivates individuals to behave more cooperatively in long term interactions.

However, the question of how costly punishment precisely affects cooperation when social identity is salient and subjects are interacting either in homogenous or heterogeneous groups, has not been examined in a one-shot public goods environment, yet.[9] Using two-person bargaining games in between-subjects design, McLeish and Oxoby (2007) provide some insight on how these institutions simultaneously influence cooperation. The authors induced group identity by letting groups jointly solve a series of questions via face-to-face communication. Next, pairs were randomly matched and informed about the other's group affiliation. After the proposer decided how much of his endowment to keep for himself and how much to allocate to the responder, in the punishment treatment the responder had the opportunity to engage in costly punishment by reducing the proposer´s income by 1.5 points for each point given. They demonstrated, first of all, that cooperation expressed by the level of offers extended both with and without punishment is higher within in-groups than in out-group matching. Secondly, they did not find significantly different cooperation levels between punishment and no-punishment treatments. This finding, however, contradicts previous research emphasizing that the threat of peer-punishment increases cooperation levels (see Fehr and Gächter 2000; 2002; Masclet et al. 2003 and Noussair and Tucker 2005). Summing up, these results provide only limited insights into two-person interactions. Evidently, they

---

[9] The interaction of social identity and third-party punishment has been the focus of a few studies already. See for instance Goette et al. (2006; 2012). In contrast to our emphasis on peer-punishment, the punisher in these studies is an individual whom the violation does not affect economically (Goette, Huffman, and Meier 2012). Additionally, Weng and Carlsson (2013) study endowment heterogeneity, peer-punishment and social identity in a work team environment using public good games. However, they do not include a treatment for identity heterogeneous groups.

cannot be easily transferred to a public good environment involving groups of more than two persons. In contrast to two-person interactions investment in punishment constitutes a second order public good in group constitutes, because every group member benefits if an offender increases his cooperation due to received sanctions. Clearly, this entails incentives to free-ride on others' costly sanctioning efforts. Consequently, these subjects form different expectations on the likelihood to be punished than subjects paired with just one person, which in turn influences their cooperativeness differently. As a result, we still lack evidence whether the hitherto reported effects of peer-punishment are valid and stable across groups, in which social identity is salient. This is, however, highly relevant since in real life most interactions are multilateral rather than bilateral.

Based on previous findings, we formulate the following research hypotheses. As extracted from the literature on public good games, peer-punishment usually tends to influence cooperation positively. Thus, we hypothesize conservatively that peer-punishment uniformly increases cooperativeness independent of whether subjects are matched with in-group or out-group members.

Hypothesis 1 (Punishment Effect Hypothesis): *The threat of peer-punishment equally impacts subjects' cooperativeness under social identity independent of group composition.*

What is more, research on social identity has shown that subjects' cooperativeness is biased to their in-group members as opposed to out-group members. Consequently, in line with our Hypothesis 1 we expect that when peer-punishment is present identity homogeneous groups tend to show higher cooperation than heterogeneous ones.

Hypothesis 2 (In-Group Bias Hypothesis): *Under punishment-threat subjects reveal higher cooperativeness when matched with in-group members than in out-group matching.*

7

III. EXPERIMENTAL DESIGN

To test the hypotheses presented above, our experiment features both within- and between-subject design elements and contains two types of treatments (see Table 1). The no-punishment (NP) treatments are based on a public good game in strategy method as introduced by Fischbacher et al. (2001). In the punishment (P) treatments we complement this public good game with monetary peer-punishment relying on a design of Noussair and Tucker (2005). [10],[11]

In total, the experiment consists of three stages (see Table 1). The participants of the experiment were assigned to groups comprising three subjects. In Stage A subjects were randomly assigned to groups (random matching) and played a one-shot public good game in strategy method without punishment – both in the P and the NP-treatments. Each group member was endowed with 20 points, which could be invested either into a public good (c_i), framed as a *project* or in their private account ($20 - c_i$). The following formula (Formula 1) denotes the payoff function without punishment for each individual i:

FORMULA 1:
Payoff function without punishment

$$\pi_i = 20 - c_i + 0.4 \sum_{k=1}^{3} c_k$$

---

[10] Please note that the NP- and P-conditions originate from two independent experiments (see Lankau et al. 2012 and Bicskei et al. 2013, respectively). The present paper merges both datasets in order to holistically analyze the effects of punishment on cooperation under social identity. The instructions to this experiment are available upon request.

[11] We are really grateful that C. Noussair and S. Tucker as well as U. Fischbacher have provided us with their experimental protocols.

Accordingly, it is more beneficial to contribute nothing to the project because every point invested returns only 0.4 points, while the private account yields exactly 1 point. Given 20 points endowment each, subjects were required to make two investment decisions, their *unconditional* and *conditional contribution* to the project. Hence, they were asked to decide how much of their endowment they would like to invest in the public good regardless of what the other group members contribute (*unconditional contribution*). Here, they were also asked for their expectations regarding the average unconditional contribution of the remaining group members. Next, subjects had to fill out a contribution table consisting of 21 entries asking them how much they *would* invest in the project if their two group members invested on average 0-20 points (*conditional contribution*).[12] Afterwards, in order to render both decisions potentially payoff-relevant, a random mechanism[13] chose one group member who had to contribute according to the conditional contribution decision. For the other two group members their unconditional contributions were thus applicable.

TABLE 1

Experimental Design

Treatments	Stage A	Stage B	Stage C
No-Punishment (NP)	Random Matching without Punishment	Induction of Social Identity	In-Group Matching (NP)
			Out-Group Matching (NP)
		-	Random Matching (NP)
Punishment (P)	Random Matching without Punishment	Induction of Social Identity	In-Group Matching (P)
			Out-Group Matching (P)
		-	Random Matching (P)

---

[12] Average investments of the other group members were rounded to the next higher integer value.
[13] One participant in each session rolled a dice to determine the respective group member. This served the purpose of making clear to the subjects that both investment decisions were to be taken seriously.

The main purpose of Stage A was to elicit subjects' initial cooperativeness independent of any identity or punishment influences, which facilitated analyses of treatment effects in later stages of the experiment. Furthermore, it helped classifying subjects based on their revealed cooperativeness into free-riders and conditional cooperators[14] (cf. Fischbacher et al. 2001; 2010; Burlando and Guala 2005) allowing for a differentiated analysis of their behavioral adjustments in later stages.

Stage B contained the group task, which served the purpose to induce social identity within the laboratory. The design we used was introduced by Ibañez and Schaffland (2012) and is based on insights from Eckel and Grossman (2005) and Chen and Li (2009).[15] The subjects were randomly reassigned to groups of different colors and they had to jointly find hidden objects in a picture and report their location given 10 minutes time. Group members were in connection via chat enabling them to communicate anonymously and discuss solutions with their group members. The answers were counted as correct only if each group member entered them correctly. This was a necessary step to foster coordination and render this task a true group exercise. The winning group was the one, which found the most objects. Since this task was intended to induce positive group experience (Eckel and Grossman 2005), only the winning team received a message of congratulation at the end of the experimental session.[16]

---

[14] Technically, free-riders are those subjects that reveal a conditional contribution of zero independent of how much their group members contribute to the public good. Subjects are classified as conditional cooperators when showing a positive and significant ($p<0.01$) spearman rank-correlation coefficient between own conditional contributions to the public good and the average contributions of their group members.

[15] We are thankful that M. Ibañez and E. Schaffland have provided us with the instructions to their identity game.

[16] In order to avoid any income effects, there were no monetary incentives for winning this game.

In the NP-treatments of Stage C subjects played a public good game similar to Stage A. In the P-treatments, after being informed how many points their group members contributed separately to the project, subjects had the possibility to distribute punishment points to each group member. [17] On the one hand, the distribution of punishment points was costly for the punisher (see Table 2). On the other hand, each punishment point distributed reduced this stage's earning of the punished by 10%. If a group member received 10 or more points his income from that stage was reduced by 100%. In order to account for possible costs of this stage of the P-treatment, and also to avoid the possibility of monetary losses, subjects received a one-off lump-sum payment of 60 points in addition to the basic endowment of 20 points (see Formula 2).

TABLE 2

Punishment Points (P_{ik}) and Cost of Punishment (K)

P_{ik}	0	1	2	3	4	5	6	7	8	9	10
$K(P_{ik})$	0	1	2	4	6	9	12	16	20	25	30

FORMULA 2

Payoff function with punishment

$$\pi_i = \left(20 - c_i + 0.4 \sum_{k=1}^{3} c_k\right) \times \frac{max\{0, 10 - \sum_{k \neq i} P_{ki}\}}{10} - \left(\sum_{k \neq i} K(P_{ik})\right)$$

$$+ \ lumpsumpayment$$

In Stage C, following a between-subject approach three different matching protocols were in effect (see Table 1). On the one hand, subjects were either matched with members of their

own identity, with whom they shared the same color when solving the group identity task in Stage B (in-group matching); or with members of different colors, thus three different identities (out-group matching). Both matching conditions are together referred to ID matchings. In order to make group identity salient, the color of an individuals' group as well as the color of her group members was clearly indicated during the game both in in-group and out-group matching. On the other hand, we carried out a control treatment both in P- and NP-treatments, in which subjects were randomly assigned to a group (random matching) and played the same one-shot public good game as in the ID matching protocols, yet without participating in Stage B (see Table 1).

The instructions were handed out successively, so that subjects learned the group assignment rule and whether they were in a P- or NP-treatment only at the relevant stage during the experiment. What is more, in the NP-treatments any information on contributions and payoffs was only revealed after the experiment's last stage. The same holds true for the P-treatments with the exception that in Stage C subjects were provided with necessary information on the contribution behavior of their group members in order to be able to make decisions about punishment points. At the end of Stage C, all subjects received detailed information on the earnings of Stage A and C and the final payment including the 2.5 € show-up fee. Moreover, in the P-treatments, subjects were only informed about the total sum of punishment points they received, yet they did not learn which group member punished them exactly.

IV. EMPIRICAL RESULTS

The experiment was conducted in the laboratory of a European university from October 2011 to July 2012 using the software z-Tree (Fischbacher 2007). We collected the decisions of 384 subjects[18] from different faculties, specifically, 177 subjects participated in the NP- and 207 subjects in the P-treatments.[19] Depending on the number of participants per session, the amount of groups and thus the number of color identities ranged from 3-8 per session. Our subjects were recruited from volunteers of the ORSEE system of the university and at the university's cafeteria. The sessions took approximately 1.5 hours and the subjects earned 15€ on average including the show-up fee.

Effects of a punishment threat on cooperativeness under social identity

At first, we investigate how the anticipation of peer-punishment affects subjects' cooperativeness in different group compositions under social identity compared to a setting without such threat (NP-treatment). Please note that although the basic decision situation of stage C was equally based on the 20 points endowment received in both P and NP-treatments, in contrast to the NP-treatment subjects in the P-treatment were given additional 60 points endowment in order to avoid bankruptcy (see section III). Consequently, differences in cooperativeness between the P and NP-treatments may be influenced by this lump-sum payment, as well. We assume, however, that this influence is equal independent of with whom subjects are matched. Given this assumption, any difference in the effect of punishment between the three matching conditions under social identity should then only be due to the

---

[18] The proportion of males was 48 percent, and the average age of the subjects was 24 years.
[19] The Appendix contains summary statistics providing details on the number of sessions and subjects per session.

effects of punishment in the respective matching. Therefore, our subsequent analyses focus on a difference-in-difference approach.

In order to keep as much within-information as possible we analyze how subjects – in their respective treatments – adjusted their conditional contributions in Stage C as opposed to Stage A. Comparing these changes between the treatments allows drawing more precise conclusions regarding differences in cooperativeness than the analysis of behavioral strategies in the particular stages independently.

To this purpose, Figure 1 depicts to what extent subjects adjusted their conditional contributions in Stage C as opposed to Stage A with and without punishment in the three matching protocols. When interacting with out-group members, punishment seems to induce the highest increase in cooperativeness amounting to almost 5 points at its peak (i.e., when the average contribution of the others is twenty). Especially, when matched with in-group members behavioral adjustments to punishment seem to be lower. Running a diff-in-diff analysis confirms that the increase in cooperativeness is significantly higher in out-group than in in-group matching (see Table 3, Model A, row 10). Comparing out-group with random matching reveals that the changes in cooperativeness are higher when interacting with out-group members. Yet, this difference is not significant. All in all, we thus have to reject Hypothesis 1. Our results clearly demonstrate that punishment has a differing effect on cooperativeness depending on group composition.[20]

---

[20] Subjects reveal similar changes in unconditional contributions (see Appendix for summary statistics).

Result 1: *Under the threat of peer-punishment subjects increase their cooperativeness significantly stronger when matched with individuals saliently belonging to different identities than with individuals of their own identity.*

FIGURE 1

The Effects of Punishment on the Average Change in Conditional Contributions
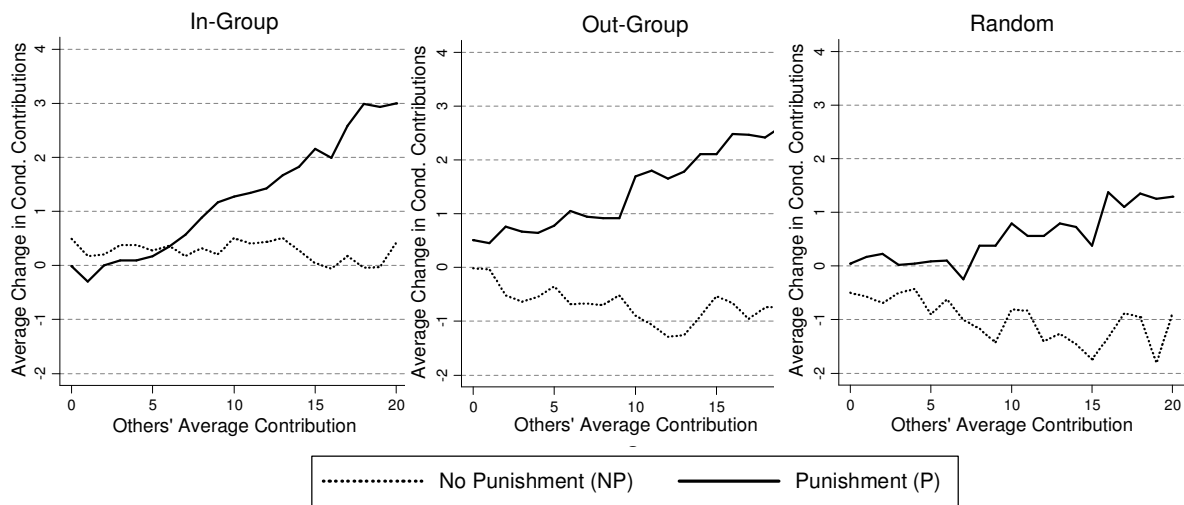
TABLE 3

The Impact of Punishment on Cooperativeness in Different Matching Conditions

Linear Regression (OLS)			Model				
			(A)	(B)	(C)	(D)	(E)
Change in Conditional Cooperation		Row	All Subjects	CCs	FRs	Others	No FRs
Punishment (P)		1	0.542*	0.475*	3.000	-0.387	0.318
			(0.312)	(0.269)	(1.870)	(0.837)	(0.271)
OUT x P		2	1.012*	0.322	3.250	0.218	0.341
			(0.523)	(0.468)	(2.254)	(0.970)	(0.434)
IN x P		3	0.705	0.287	1.056	1.400	0.506
			(0.460)	(0.463)	(2.240)	(1.001)	(0.416)
No Punishment (NoP)		4	-1.008*	-2.193***	1.177	-0.179	-1.445**
			(0.540)	(0.720)	(1.103)	(0.888)	(0.583)
OUT x NoP		5	0.317	1.468*	-1.177	-0.959	0.631
			(0.651)	(0.831)	(1.103)	(1.567)	(0.722)
IN x NoP		6	1.274**	2.365***	0.537	-0.207	1.466**
			(0.620)	(0.825)	(1.494)	(0.919)	(0.656)
IN: P - NoP		7	0.981*	0,589	2,342	1.400*	0.802*
OUT: P-NoP		8	2.244***	1.522***	6.25***	0,97	1.473***
Control: P-NoP		9	1.550*	2.668***	1,823	-0,207	1.764***
Diff-in-Diff Analyses	OUT vs IN	10	1.264*	0,933	3.908*	-0.431	0.670
	OUT vs. Control	11	0,695	-1,145	4.427*	-1,177	-0.290
	IN vs. Control	12	-0,569	-2.078**	0,519	1,608	-0.960
Punishment	IN - OUT	13	-0.306	-0.035	-2.194	1,183	0,165
			(0.540)	(0.538)	(1.767)	0,736	0,462
No Punishment	IN - OUT	14	0.957**	0.898	1.714*	0,752	0,835
			(0.474)	(0.578)	(1.007)	(1.312)	(0.523)
Number of obs			8.064	5.334	1.134	1.596	6.930
R-squared			0.058	0.060	0.390	0.022	0.038

Note: standard errors clustered over individuals, Stat. Sign.: *p<0.1, **p<0.05, ***p<0.01
Coefficients as off row 7 are calculated using post-regression tests for linear combinations

Next, we ask how certain cooperation types with differing social attitudes adjust their cooperativeness. To do so, we classify subjects – based on their initially revealed cooperativeness in Stage A of the experiment – into free-riders (FR) and conditional cooperators (CC). While free-riders reveal purely selfish cooperation preferences, conditional cooperators tend to be motivated by a social preference based on positive reciprocity (cf. Falk 2003; Lankau et al. 2012). Consequently, the more their group members contribute to the public good, the higher their own contribution is in return. Yet, even conditional cooperators

do not perfectly reciprocate their group members' contributions and exhibit a certain self-serving bias by frequently contributing less to the public good than their group members (Fischbacher, Gächter, and Fehr 2001; Lankau, Bicskei, and Bizer 2012). Subjects, which fell into neither category, are termed "others".[21] Table 4 summarizes the distribution of those cooperation types.

TABLE 4

Frequency and Distribution of Cooperation Types in Stage A

Cooperation Type	Frequency	Percent
Free-Rider (FR)	54	14.06
Conditional Cooperators (CC)	254	66.15
Other	76	19.79
Total	**384**	**100**

Figure 2 and 3 show the impact of peer-punishment on free-riders' and conditional cooperators' cooperativeness in in-group, out-group and control treatments, respectively. We find, on the one hand, that independent of how free-riders are matched, peer-punishment increases their cooperativeness. The strongest behavioral adjustment is present when free-riders interact with out-group members, which is significant at the 10 percent level (see Table 3, Model C, row 10). While they are not cooperative at all without the threat of peer-punishment, under punishment they are for instance willing to contribute 12 points on average to the public good if their group members contribute 20. When matched with in-group members or in random matching these differences are lower and merely amount to approximately 4 points on average. Indeed, the linear regression in Table 3 (Model C, row 10

and 11) confirms that the impact of punishment on free-riders' cooperativeness in out-group matching is significantly stronger than in the in-group or control treatment.

On the other hand, the punishment effect on conditional cooperators does not seem to be as pronounced as in case of free-riders (see Figure 3 and Table 3, Model B and C, row 7, 8). Moreover, we find that conditional cooperators reveal a very similar behavioral adjustment under punishment when matched with in-group and with out-group members (Table 3, Model B, row 10).

Result 2: *In ID matchings free-riders are prone to show stronger adjustments in cooperativeness in response to peer-punishment than conditional cooperators. It is strongest when matched with out-group members.*

In order to explain this differential reaction to punishment, we argue, on the one hand, that socially oriented conditional cooperators voluntarily contribute to the public good so that the threat of punishment does not evoke a particularly strong increase in their cooperativeness. Purely self-interested free-riders, on the other hand, strongly adjust their cooperation strategy under punishment due to the perceived threat of monetary losses following their uncooperative behavior. This anticipation of peer-sanctioning seems to be especially strong in groups with individuals of differing identities.

FIGURE 2

The Effects of Punishment on Average Changes in Conditional Contribution of Free-Riders
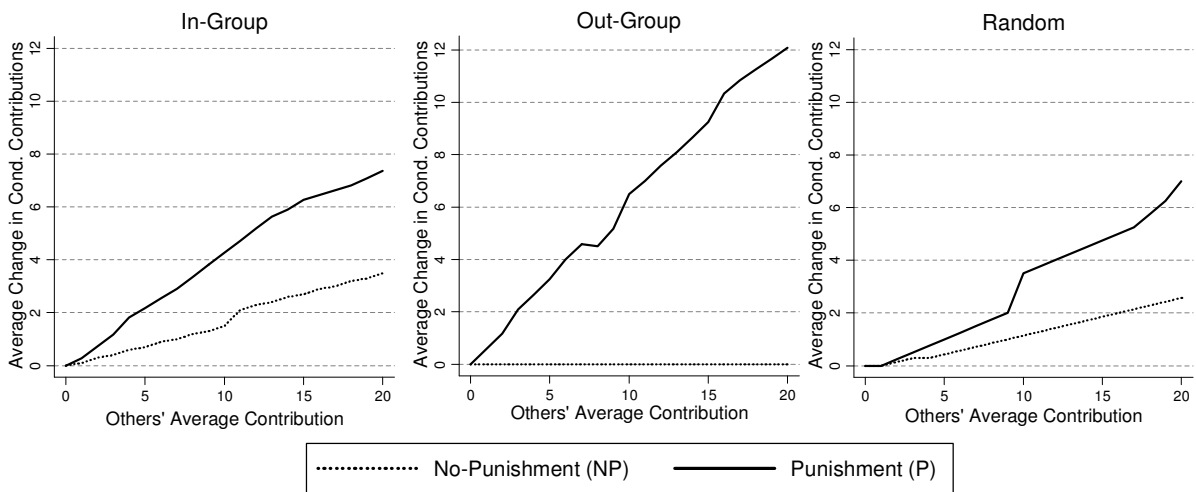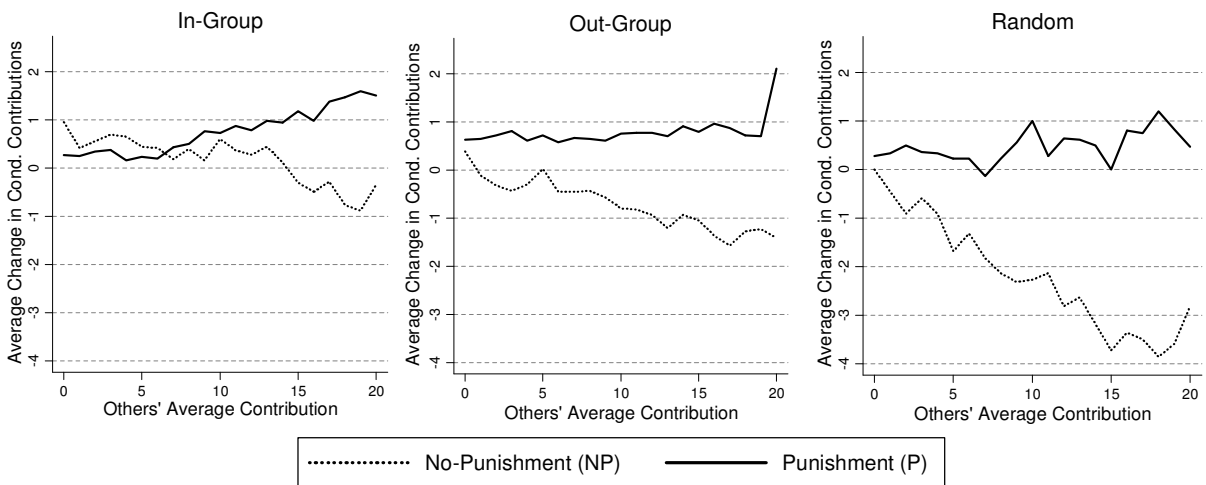

FIGURE 3

The Effects of Punishment on Average Changes in Conditional Contribution of Conditional
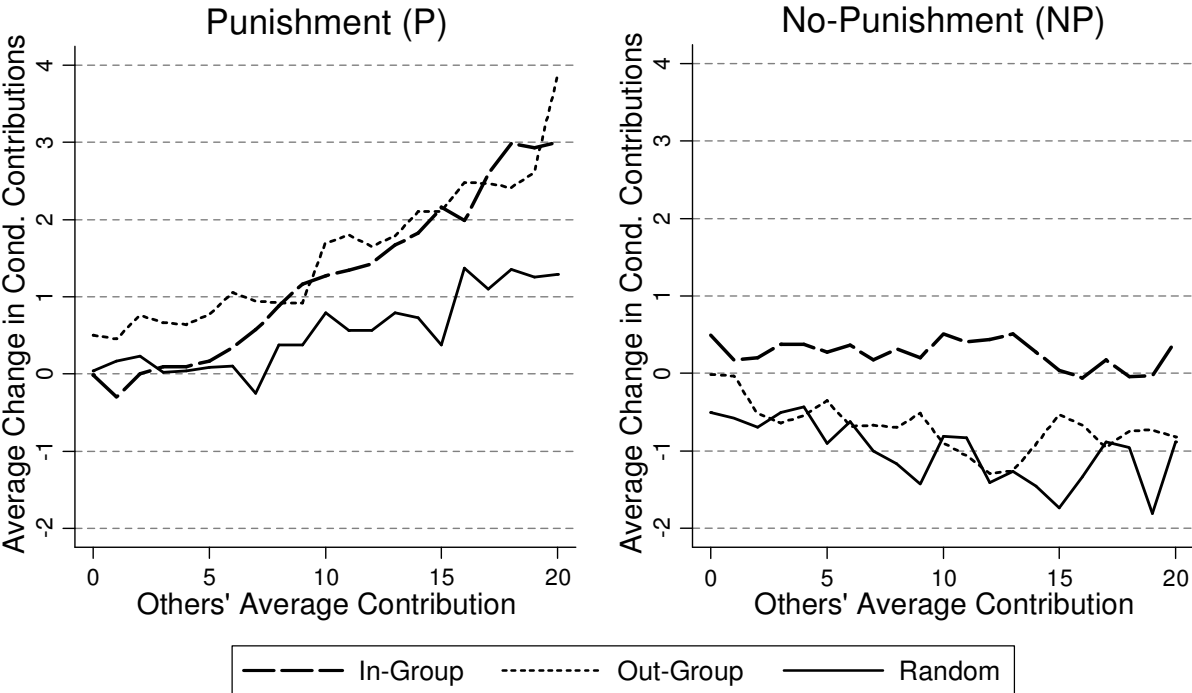
Cooperators

Effects of Social Identity under Peer-Punishment-Threat

After having established that peer punishment drives cooperativeness particularly in identity heterogeneous groups, we proceed by analyzing whether subject's cooperativeness differs dependent on group composition when a punishment threat is present. Figure 4 depicts the

19

average change in conditional cooperation between Stage C and A under punishment as influenced by the three group assignment conditions (left panel) contrasted by the corresponding cooperativeness without punishment (right panel).

FIGURE 4

ID Matchings Effects on Changes in Conditional Cooperation with and without Punishment

Firstly, under punishment subjects seem to be similarly cooperative in in- and out-group matching (left panel). For low levels of others' contribution to the public good they even contribute on average slightly more when matched with out-group members. On the whole, however, we do not find any significant difference in conditional cooperation between both ID treatments under punishment (see Table 3, Model A, row 13). This starkly contrasts the findings on subjects' cooperativeness without punishment (see Table 3, Model A, row 13). In line with prior research on the effects of social identity on cooperation (cf. Chen and Li 2009;

Eckel and Grossman 2005), subjects' cooperativeness without punishment is significantly biased towards their in-group members (right panel, for regression result see Table 3, Model A, row 14).[22] Consequently, we argue that the fear created by a punishment institution overrides the reluctance to cooperate with out-group members and renders individuals to be as cooperative as in in-group matching. Evidently, as summarized by Result 1, behavioral adjustments through punishment seem to be comparatively strongest when subjects are matched with out-group members. This is most likely due to an anticipation of higher punishment by individuals that do not belong to their own identity, which seems to be mainly driven by free-riders. They exhibit the highest increase in cooperativeness under punishment when matched with out-group members (Result 2). Indeed, when dropping free-riders from the regression analysis in Table 3 (Model E, row 10), we no longer find a significant difference in punishments' impact on subjects' cooperativeness in in-group and out-group matching.

Mild support for the claim that subjects fear comparably strong punishment by outsiders in response to antisocial behavior is given by considering their actual punishment behavior (see summary statistics in the Appendix). Subjects in out-group matching distributed on average 2.34 punishment points to group members that contributed less than they did (non-spiteful punishment). In in-group and random matching non-spiteful punishment merely amounted to 1.51 and 2.25 points, respectively.[23] Consequently, since members of identity heterogeneous groups tend to punish deviant behavior on average more severely than identity homogenous

[22] The induction of social identity in Stage B was identical both in the NP as well as the P treatments. Consequently, the existence of an in-group bias without punishment clearly shows that the non-existence of such bias with punishment cannot be explained by an unsuccessful identity induction.

[23] Wilcoxon Rank Sum Tests: In vs. Out, $p<0.05$; Cont vs. In, $p<0.01$; Out vs. Cont, $p>0.1$. For a detailed discussion of subjects' punishment behavior and its relation to anger-like emotions see Bicskei et al. (2013).

groups or randomly matched groups, it is not unreasonable to assume that subjects anticipate such reaction in response to unsocial behavior and adjust their cooperativeness accordingly.[24]

Overall, since we no longer observe an in-group bias, we have to reject Hypothesis 2, as well. This result is inconsistent with McLeish and Oxoby (2007) who find in two-person interactions that both with and without punishment cooperation is higher in in-groups than in out-groups. However, they did not find any punishment effect on cooperation either.

Result 3: *The presence of peer-punishment eliminates the existence of an in-group bias, which is prevalent when social identity is salient and no punishment is available.*

Lastly, subjects' cooperativeness under punishment both in in- and out-group matching is slightly elevated compared to random matching of the control treatment in which no social identity is induced (Figure 4, left panel). In in-group matching this difference is merely borderline significant (Table 3, Model A, row 3, p=0.127). In out-group matching significance is given at the 10 percent level (Table 3, Model A, row 2). Together, these results suggest that a simultaneous application of peer-punishment and social identity is a viable option to boost cooperativeness in public goods provision.[25]

---

[24] Please note that due to our experimental design, the actual amount of punishment points received cannot influence subjects' cooperativeness. Only after they make their contribution decisions do subjects learn how much they have been punished. It is thus impossible to directly conclude from punishment points to cooperativeness. It's much rather the anticipation of punishment that drives subjects' behavior.

[25] One interesting reflection regarding our research design, pointed out to us by a commentator, is the fact that both ID-matching protocols entail a higher cooperativeness than the random matching protocol under punishment could be simply due the design of the control treatment. Specifically, it did not require subjects to take part in Stage B, so that missing the experience of solving the picture task explains the observed difference. Yet, if this was true we should observe similar differences between those treatments without punishment, as well. However, without punishment subjects in out-group matching reveal a similar cooperativeness as those of the control treatment (see Figure 4 and Table 3).

Result 4: *Under punishment subjects reveal an increased cooperativeness in in- and out-group matching compared to groups without saliency of social identity.*

On the one hand, this result complements to some degree the findings by Noussair and Tucker (2005) who established that when peer-punishment is combined with another institution, namely with the possibility of non-monetary social sanctions, it results in a higher cooperativeness than when applied by itself. On the other hand, it contradicts Weng and Carlsson (2013), who find that punishment has a similar effect on cooperativeness in groups without a salient social identity and in groups sharing a common salient identity.

V. CONCLUSION

In this article, we investigate the question of how the threat of peer-punishment and social identity simultaneously affect cooperation in the provision of public goods. In particular, our aim is to provide insights on the influence of punishment on subjects' cooperativeness in homogeneous groups bound together by a common social identity and in heterogeneous groups consisting of individuals saliently belonging to different identities. To this purpose, we induce social identity in the laboratory and devise one-shot public good games using the strategy method both with and without the institution of costly peer-punishment. Our experiment is the first that asks for effects of peer-punishment in public goods provision dependent on whether subjects interact in identity-homogenous or identity-heterogeneous groups. In particular, it adds to McLeish and Oxoby (2007), who focus on two-person interactions, by incorporating the influence of a group setting on subjects' anticipated

---

Consequently, it is not the participation in Stage B but rather the identification created by it and punishment, which explain the observed effects.

punishment and their cooperation under social identity. Punishment within groups itself constitutes a second order public good that incentivizes subjects to free-ride on group members punishment efforts.

The results of our study indicate first of all that group composition plays a crucial role in how a peer-punishment threat affects cooperation under social identity. The strongest increase in cooperativeness can be observed among subjects in out-group matching, especially for those who were classified as free-riders based on their initially revealed cooperativeness. This is most likely due to an anticipation of comparably strong punishment by individuals who do not belong to their own identity. Secondly, in contrast to the institutional setting without peer-punishment, in which individuals' cooperativeness is clearly biased towards members of their own identity (in-group bias), the anticipation of stronger peer-punishment by out-group members generally seems to erode this reluctance to cooperate with out-group members and renders individuals as cooperative as in in-group matching. Lastly, our results indicate that social identity and peer-punishment complement each other. In in- and out-group matching under punishment subjects' cooperativeness is elevated compared to the control treatment, in which group members were randomly matched and no social identity was induced. Concluding, the combination of social identity with a peer-punishment institution seems to be a viable option to foster subjects' cooperativeness compared to situations when social identity is not salient.

Summing up, it is empirically established that fragmentation and polarization of societies restrain the provision of public goods (Eaton, Eswaran, and Oxoby 2011) since individuals strongly identifying with particular groups are willing to cooperate less with members of different identities. We established, however, that implementing peer-punishment erodes the

reluctance to cooperate in identity heterogeneous groups and renders people as cooperative as when being part of an identity homogenous group.

Summary Statistics

	NP-Treatment			P-Treatment		
	IN	OUT	CONT	IN	OUT	CONT
n	69	66	42	84	75	48
Session #	3	3	3	7	6	3
Unconditional Contribution- Stage A	5.83	7.38	8.29	7.37	6.33	9.33
	(0.72)	(0.74)	(0.94)	(0.66)	(0.66)	(0.98)
Contribution (real)- Stage A	4.10	7.18	8.26	6.70	5.76	7.77
	(0.60)	(0.74)	(0.98)	(0.65)	(0.64)	(0.96)
Change in unconditional Contributions	0.99	-1.06	-0.38	2.37	3.68	0.94
(Stage C - Stage A)	(0.64)	(0.60)	(0.46)	(0.50)	(0.72)	(0.69)
Change in real Contributions	1.22	-1.73	-1.52	2.55	3.61	1.21
(Stage C - Stage A)	(0.71)	(0.72)	(0.88)	(0.64)	(0.71)	(0.89)
Total Punishment received- Stage C	n/a	n/a	n/a	2.14	2.60	2.83
				(0.35)	(0.38)	(0.43)
Total Punishment distributed- Stage C	n/a	n/a	n/a	1.07	1.30	1.42
				(0.15)	(0.18)	(0.22)
Non-spiteful Punishment distributed- Stage C	n/a	n/a	n/a	1.51	2.34	2.25
				(0.25)	(0.32)	(0.30)

Note: Standard errors in parantheses.

References

Anderson, Christopher M., and Louis Putterman. 2006. "Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism." *Games and Economic Behavior* 54 (1) 1-24.

Bicskei, Marianna, Matthias Lankau and Kilian Bizer. 2013. *Negative Reciprocity and Anger-Like Emotions in Homogenous and Heterogeneous Groups*. Unpublished Manuscript.

Blackwell, Calvin, and Michael McKee. 2003. "Only for my own neighborhood? Preferences and voluntary provision of local and global public goods "*Journal of Economic Behavior & Organization* 52 (1) 115-131.

Bochet, Olivier, Talbot Page, and Louis Putterman. 2006. "Communication and punishment in voluntary contribution experiments." *Journal of Economic Behavior & Organization* 60 (1): 11–26.

Burlando, Roberto, and Francesco Guala. 2005. "Heterogeneous Agents in Public Goods Experiments." *Experimental Economics* 8 (1): 35–54.

Chakravarti, Surajeet, and Miguel A. Fonseca. 2012. "The effect of social fragmentation on public good provision: an experimental study" University of Exeter, Economics Department, Discussion Paper Series, Paper No. 12/07.

Chakravarti, Surajeet, and Miguel A. Fonseca. 2012. "Discrimination via Exclusion: An Experiment on Group Identity and Club Goods". University of Exeter, Economics Department, Discussion Paper Series, Paper No. 13/02.

Charness, G., L. Rigotti, and A. Rustichini. 2007. "Individual behavior and group membership." *The American Economic Review* 97 (4): 1340–52.

Chen, Roy, and Yan Chen. 2011. "The Potential of Social Identity for Equilibrium Selection." *American Economic Review* 101 (6): 2562–89.

Chen, Yan, and Sherry X. Li. 2009. "Group Identity and Social Preferences." *The American Economic Review* 99 (1): 431–57.

Eaton, B. C., Mukesh Eswaran, and Robert J. Oxoby. 2011. "Us and Them: the origin of identity, and its economic implications." *Canadian Journal of Economics* 44 (3): 719–48.

Eckel, C. C., and P. J. Grossman. 2005. "Managing diversity by creating team identity." *Journal of Economic Behavior & Organization* 58 (3): 371–92.

Falk, Armin. 2003. "Homo Oeconomicus versus Homo Reciprocans: Ansätze für ein neues Wirtschaftspolitisches Leitbild?" *Perspektiven der Wirtschaftspolitik* 4 (1): 141–72.

Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *The American Economic Review* 90 (4): 980–94.

———. 2002. "Altruistic punishment in humans." *Nature* 415 (6868): 137–40. doi: 10.1038/415137a.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10: 171–78.

Fischbacher, Urs, and Simon Gächter. 2010. "Social preferences, beliefs, and the dynamics of free riding in public goods experiments." *The American Economic Review* 100 (1): 541–56.

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

Goette, Lorenz, David Huffman, and Stephan Meier. 2006. "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups." *American Economic Review* 96 (2): 212–16.

———. 2012. "The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups." *American Economic Journal: Microeconomics* 4 (1): 101–15.

Harris, Donna, Benedikt Herrmann, and Andreas Kontoleon. 2012. "When to Favour Your Own Group? The Threats of Costly Punishments and In-group Favouritism." University of Exeter, Economics Department, Discussion Paper Series, Paper No. 628.

Hoff, Karla, and Priyanka Pandey. 2006." Discrimination, Social Identity, and Durable Inequalities." *American Economic Review* 96 (2): 206–211.

Ibañez Diaz, Marcela, and Elke R. Schaffland. 2012. *Group identity and leadership: An experimental analysis.* Unpublished Manuscript.

Lankau, Matthias, Marianna Bicskei, and Kilian Bizer. 2012. "Cooperation preferences in the provision of public goods: An experimental study on the effects of social identity." Cege Discussion Papers 148. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2191469.

Masclet, David, Charles Noussair, Steven Tucker, and Marie-Claire Villeval. 2003. "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism." *American Economic Review* 93 (1): 366–80.

McLeish, K. N., and R. J. Oxoby. 2007. "Identity, cooperation, and punishment." IZA Discussion Paper 2572. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=961379.

McLeish, Kendra N., and Robert J. Oxoby. 2011. "Social interactions and the salience of social identity." *Journal of Economic Psychology* 32 (1): 172–78.

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

Noussair, Charles, and Steven Tucker. 2005. "Combining monetary and social sanctions to ptomote cooperation." *Economic Inquiry* 43 (3): 649–60.

Page, Talbot, Louis Putterman, and Bulent Unel. 2005. "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency." *Economic Journal* 115 (506): 1032–53.

Sefton, Martin, Robert Shupp, and James M. Walker. 2007. "The Effect of Rewards and Sanctions in Provision of Public Goods." *Economic Inquiry* 45 (4): 671–90.

Selten, Reinhard. 1967. "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments." In *Beiträge zur experimentellen Wirtschaftsforschung*. Vol. 1. Edited by H. Sauerland, 136–68. Tübingen: Mohr.

Tajfel, Henri, and John Turner. 1979. "An Integrative Theory of Intergroup Conflict." In *The Psychology of Intergroup Relations*. Edited by Stephen Worchel and William Austin, 33–47. Monterey: CA: Brooks/Cole.

———. 1986. "The Social Identity Theory of Intergroup Behaviour." In *The Psychology of Intergroup Relations*. Edited by Stephen Worchel and William Austin, 7–24. Chicago: Nelson-Hall.

Weng, Qian and Fredrik Carlsson. 2013. "Cooperation in teams: the role of identity, punishment and endowment distribution." University of Gothenburg, Department of Economics, Working Papers in Economics, Paper No. 551.