JENA ECONOMIC RESEARCH PAPERS

2010 – 066

Group Reciprocity

by

David Hugh-Jones
Martin A. Leroch

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

Jena Economic Research Papers 2010 - 066

Group Reciprocity

*David Hugh-Jones and Martin A. Leroch**

24 September 2010

Abstract

People exhibit *group reciprocity* when they retaliate, not against the person who harmed them, but against somebody else in that person's group. Group reciprocity may be a key motivation behind intergroup conflict. We investigated group reciprocity in a laboratory experiment. After a group identity manipulation, subjects played a Prisoner's Dilemma with others from different groups. Subjects then allocated money between themselves and others, learning the group of the others. Subjects who knew that their partner in the Prisoner's Dilemma had defected became relatively less generous to people from the partner's group, compared to a third group. We use our experiment to develop hypotheses about group reciprocity and its correlates.

Keywords: reciprocity, groups, conflict

JEL classification: D74, C92

* David Hugh-Jones: Max Planck Institute of Economics, Kahlaische Strasse 10, 07745 Jena, hugh-jones@econ.mpg.de. Martin Leroch: University of Hamburg, leroch@econ.uni-hamburg.de. We would like to thank Vittoria Levatti, Eva Steiger, Johannes Weisser, Ro'i Zultan, Werner Gueth, David Reinstein, Ryan McKay, Eva van den Broek and seminar participants at the MPI; and the Max Planck Institute ESI group hiwis and administrative staff: Martin Beck, Nadine Erdmann, Adrian Liebtrau, Christian Williges, Christian Streubel, Claudia Zellmann and especially Christoph Goering.

> Saturday 27th January 2001 – A 23 year old white man, Mark Clayton, was stabbed and assaulted by a group of Asian teenagers in a subway near Manchester Street, Oldham, in what remains the worst racist attack of the current troubled period. ...
>
> February/March – Some continuing racist violence, e.g. an attack on an Asian man outside a pub by four white men on 23rd February ....
>
> April – A series of racist attacks, mainly by groups of Asian youths on lone white men, continues to make the headlines in Oldham....
>
> 21st-25th May – A series of incidents occurred during this week at Breeze Hill School.... On 21st May, a group of white ex-pupils attacked some Asian pupils, and there was stone throwing and verbal abuse. ...
>
> Saturday, 26th May – This day was the first, and most serious day, of the Oldham riots.
>
> Ritchie (2001)

Introduction

Laboratory experiments have persuaded many economists that humans are reciprocators, who will pay so as to punish unkind actions, both towards them and others. Real-world examples like the above show that people sometimes strike back, not against the individual who caused the harm, but against other people in that individual's group. We call this kind of retaliation *group reciprocity*.

Acts of group reciprocity are central to human conflict. Almost all violent conflicts involve calls for revenge against an opposing group for previous harms, which may be real or imaginary, and recent or very old.[1] Group reciprocity can also be seen in ordinary economic behaviour. For instance, consumers may boycott the produce of countries whose policies they disapprove of. German products were boycotted in Greece as reaction to a German news magazine's headline calling Greeks “Crooks in the European Family” in early 2010.[2] Companies and teams can also be relevant groups for reciprocity: Dick Fuld, former CEO of Lehman Brothers, advised his employees “if one of your teammates gets attacked, fight back like hell” (Sorkin 2009)!

Although group reciprocal *behaviour* has been attested beyond reasonable doubt, that does not prove that humans have a psychological disposition to it. Acts of group reciprocity could be undertaken by materially self-interested, rational actors. For example, doing so might help someone to build a reputation for toughness. Or, not doing so might expose one to punishment from others in one's own group: one eyewitness of the Rwandan genocide stated “ten percent helped; 30 percent were forced to kill; 20 percent killed reluctantly; 40 percent killed enthusiastically” (Mamdani 2001).

However, as the quote suggests, some people may have a psychological disposition to group-reciprocate. To investigate whether this is the case, we run a laboratory experiment in which

1 For many examples, see Horowitz (1985; 2001).

2 The boycott of a company’s products may eventually also turn into violence. In 1995, for example, the Shell Corporation decided to sink the oil drilling platform “Brent Spar” in the Northern Sea, a decision which triggered massive protest all over Europe. Further, as German newspapers reported, several tenants o Shell Patrol Stations in Gemany received bomb threats, at least one of them actually receiving a letter bomb.

there is no future interaction, and therefore no strategic incentive to group-reciprocate.[3] We find that even in these circumstances, some people do group-reciprocate.

Acts of group reciprocity may be carried out by either the attacked individual, or his or her group members. For instance, most participants in the Oldham race riots had probably not themselves been victims of violence by the other ethnic group. Instead, they may have seen themselves as taking action on behalf of their fellow group members. In this paper, we investigate direct group reciprocity, i.e. reciprocity by the harmed individual himself, leaving indirect group reciprocity for future work. Reciprocity can normally be positive as well as negative. That is, one may reward others for their helpful behaviour as well as punishing them for harmful behaviour. Some of our results could be framed in terms of either positive or negative reciprocity. We mainly describe them in terms of negative reciprocity, because we believe that this is more important and salient in the real world.

In the next section we review the relevant literature in psychology and economics. We then describe our experiment and results. Lastly we draw conclusions.

Related Literature

Experimental economists have recently become interested in the effects of group membership on behaviour, a topic that has long fascinated social psychologists. Hargreaves Heap and Zizzo (2009) show experimentally that people will put a monetary value on group membership. Further, people discriminate between in- and outgroups. They cooperate more

3 Like all economic experiments on adults, our experiment cannot determine whether psychological dispositions to group reciprocity are innate, or are learned from repeated exposure to situations where group reciprocity was appropriate behaviour.

with in-group members (see, e.g., de Cramer and van Vugt 1999 and Guala et al. 2009); give in-group members more, reward them more for good behaviour, and punish them less (Chen and Li 2009). Debate continues over whether this different behaviour is driven by expectations of future reciprocity (Yamagishi and Kiyonari 2000; Yamagishi and Mifune 2008, 2009) or by pure preferences (Guala et al. 2009).As psychologists have long noted (Brewer 1999), the existence of intergroup bias per se does not imply that people will behave unkindly towards members of an outgroup, only that they will behave more kindly towards members of their ingroup. For example, when asked to allocate a "bad" (exposure to aversive noise) among ingroup and outgroup members, subjects showed no ingroup bias (Mummendey 1992). Yet, in real-world conflicts, real groups often do terrible things to each other. Furthermore, some lab experiments have induced harmful conflicts between groups. For instance, Bornstein (1992, 2003) observed that subjects contributed to a public good which benefitted members of the ingroup and harmed outgroup members (and was inefficient for the "society" as a whole). Thus, subjects were willing to harm outgroups so as to help the ingroup. Even more strikingly, Abbink and Herrmann (2009) created a ten-round vendetta game, with subjects in two groups: in each round, subjects could deduct money from the other group, at a cost to themselves. The low observed rates of this harmful behaviour (13%) were tripled (40%) by the addition of a symbolic reward.[4]

Therefore, there seems to be a gap between individual-level motivation and collective outcomes, whether in the lab or the field. We believe that dispositions to group reciprocity may be part of the answer. Individual biases towards one's own group, which are (relatively) innocuous in ordinary circumstances, can be transformed into something much more

4 But not every experiment reproduces such effects; see e.g. Halevy et al. (2008).

dangerous when one perceives a threat or insult from another group.[5] Some analyses of field data support this idea. Shayo and Zussman (2010) show that Israeli judges were more likely to find in favour of Israeli plaintiffs against Arab defendants shortly after terrorist incidents. This could be explained by increased solidarity with the ingroup, or by bias against an outgroup which is perceived as responsible for an attack. (Though Arab judges were also more likely to find in favour of Arabs against Israelis after an attack, which favours the ingroup solidarity interpretation.) In a military context, Kocher et al. (2008) show that indiscriminate US bombing in Vietnam increased Vietcong control, suggesting that it drove Vietnamese civilians "into the arms of the rebels".

Psychologists have investigated third-party group reciprocity under the term "vicarious retribution" (Lindgren et al. 2006). However, while experiments have shown that subjects will verbally express a demand for sanctions in retaliation for a presumed intergroup insult (Stenstrom et al. 2008), we know of no psychological experiment demonstrating intergroup retribution with real material consequences. Some economic experiments are suggestive: for instance, in Abbink and Herrmann (2009), rates of harmful actions were increased by the other group's harmful actions in the previous round. Similarly, Reuben and van Winden (2008) show that friends are more likely to coordinate their punishment of an unkind proposer, and to punish more severely, in a power-to-take game. These experiments cannot wholly rule out strategic motivations, however – whether within-game in the multiple round setup, or outside the lab for friends.Our experiment therefore tests whether people will reciprocate the actions of one group member when they make choices affecting other

5 Moreno (2008) provides a game-theoretic analysis of group fairness, which includes the idea of intergroup reciprocity.

members of that group. We rule out strategic motivations, including expectations about other players' behaviour towards oneself.

Experimental design

Our experiment was run at the Jena laboratory of the Max Planck Institute for Economics, using zTree.[6] 180 subjects, recruited using ORSEE, took part during 6 sessions. Each session lasted about 1 hour.

In each session of our experiment, 30 subjects were randomly allocated into three groups named *red*, *green* and *blue*. In order to increase group entitativity (the sense that each group forms a single entity) and identification, the groups played a team game. This was an adapted version of the card game Pelmanism, in which players have to find pairs of identical cards from a face-down deck. In our computer version, the whole group voted for which card to turnover, and one selected member then observed the votes and chose a card. The group which found the most pairs was awarded 5 ECU. However, the winning group was not revealed until the end of the experiment.

Each subject was given a player number, running from 1-10 within each group. These numbers, together with the group names, were used to identify subjects to each other throughout the remainder of the experiment.

After the Pelmanism game, subjects were allocated into pairs, each pair containing subjects from two different groups. They were informed of their partner's group and player number. They then played a one-shot prisoner's dilemma (PD). Afterwards, subjects were shown their

6 Screenshots of the interface are available online in the file screenshots.zip

partner's choices and the resulting payoffs for both players. We will call each subject's PD partner's group the Partner group; the remaining group is the Other group. For instance, if a red group member played the prisoner's dilemma against a green group member, then for her the green group is the Partner group, and the blue group is the Other group.

Lastly, each subject made a set of ten binary choices, allocating money between him- or herself and other subjects, always identified by their player number and group. Four allocations were between the subject and a member of the Partner group. Four further allocations involved identical amounts of money as the first four, but were between the subject and a member of the Other group. Two allocations were between a Partner group member and an Other group member, with the subject herself not being involved. Subjects never made allocations involving their previous PD partner – only involving other members of the same group. Subjects could easily observe this, because the PD partner's group and player number, along with the outcome of the PD, was shown onscreen throughout the 10 allocations.

The order of the allocations was randomized. At the end of the experiment a single allocation was randomly chosen, and the relevant payoffs were implemented.

After the experiment, subjects were given a questionnaire including demographics, measures of group identity, and debriefing questions. The winning group of the Pelmanism game was then revealed. Finally, subjects were called up and paid privately for their winnings from Pelmanism game, the Prisoner's Dilemma, and the randomly chosen allocations (their own and others'), as well as a EUR 2.50 showup fee.

Empirics

The ten allocation choice sets are shown in Table 1. Subjects chose either option A or option B: the amounts allocated are shown in the corresponding columns.[7] We number the five different kinds of choices from 1-5. Each choice was made twice, against a member of the Partner group and against a member of the Other group (or for choice set 5, the positions of Partner and Other players were switched). The two choices within each pair are differentiated by the suffix P or O.

Each set had one option which gave strictly less to the other player. We call this the *unkind* option. Each also had one *selfish* option which gave strictly more to the subject herself; the other option was *costly*. Lastly one option was *equal* because it had a strictly lower difference between the subject and the other player.

7 On subjects' screens, the order and labels of the columns were randomized.

1P	A	B
s	2	1
P	5	2

1O	A	B
s	2	1
O	5	2

2P	A	B
s	4	3
P	3	0

2O	A	B
s	4	3
O	3	0

3P	A	B
s	1	2
P	4	1

3O	A	B
s	1	2
O	4	1

4P	A	B
s	4	5
P	3	0

4O	A	B
s	4	5
O	3	0

5P	A	B
O	3	1
P	1	2

5O	A	B
P	3	1
O	1	2

S = self, P = prisoner's dilemma PARTNER's group, O = OTHER group
Numbers are ECUs.
Choices 1-4 are self-other choices. Choice 5 allocates money between two other subjects.
In the self-other choices:
Choice B is always unkind.
In choices 1-2, the selfish choice is kind.
In choices 3-4, the selfish choice is unkind.
In choices 2 and 4, the equal choice is kind.
In choices 1 and 3, the equal choice is unkind.

Table 1: Choice sets

In four of the 8 choices, the unkind option was also selfish. Thus, choosing the unkind option can reflect simple material self-interest, while the kind option may be altruistic. In the other four choices, the unkind option was not selfish: to harm the other player, subjects had to harm themselves. These choices can be thought of as “costly punishment”. There are real world analogues in conflict situations, in which people may put themselves at risk in order to harm those in other groups.

Because subjects had no interaction with members of other groups until the Prisoner's Dilemma, the choice of their partner to *cooperate* or to *defect* was random and uncorrelated with any of the subject's characteristics. Therefore, we can treat partners' choices as an exogenous between-subjects treatment. We can then examine subjects' levels of unkindness towards Partner group members (members of the PD partner's group) and towards Other group members, comparing these differences among subjects whose partners defected, and among those whose partners cooperated. We can also repeat this analysis for particular choices, including those where unkindness is selfish or not, and those where the unkind choice is equal or not; and for particular subject groups, including those who did or did not themselves cooperate in the Prisoner's Dilemma.

Results

Descriptive and demographic statistics are shown in Table 9 in the Appendix. In particular, we deliberately invited a high proportion of males. Since a large majority of active participants in violent group conflicts are male (see for instance Archer 2004 or Scheff 2003), we believed that group reciprocity might be easier to observe in a male-dominated environment.

In the Prisoner's Dilemma, 64 subjects defected and 116 cooperated. The above-average cooperation rates are probably due to framing: in the PD description, we described each player as having an “endowment” of 2 ECUs, some of which would be “lost” if the other player defected.

We first examine the self-other choice sets (numbers 1-4). Before differentiating between decisions for Partner and Other groups, we look at decisions in general.

		Subject's PD action		
Choice sets	Unkind decision is...	… cooperated	… defected	P-value (χ^2 test)
1	Costly, equal	10.7%	14.8%	0.337
2	Costly, unequal	6.0%	2.3%	0.187
3	Selfish, equal	85.8%	88.3%	0.612
4	Selfish, unequal	50.4%	63.3%	0.025

Table 2: Unkind decisions by subject's action in the Prisoner's Dilemma

Table 2 shows the percentage of unkind decisions for each of the self-other choice sets, split by whether the subject cooperated or defected in the Prisoner's Dilemma. Except where the selfish decision was both harmful to the other and inegalitarian, large majorities chose the selfish decision, and there was no significant difference between those who cooperated or not in the PD. This was true even though the selfish decision in choice sets 3 & 8 was inefficient (lowered the total amount paid to the pair). However, when the selfish decision was harmful to the other and inegalitarian (and also inefficient), decisions were more evenly split, and significantly more PD defectors were selfish than PD cooperators.

Table 3 shows the percentage of unkind decisions, split by whether the subject's partner in the Prisoner's Dilemma cooperated or defected. The last column reports P-values from a χ^2 test. Costly unkindness in general seems to have been affected by the PD partner's choice, but in opposite ways depending on whether the unkind decision increased or decreased inequality: players became more likely to choose the more equal option, whether this harmed or helped the other player. Possibly, the other player's cooperation in the PD activated a norm of cooperative behaviour, which had an egalitarian component.

		PD partner...		
Choice sets	Unkind decision is...	… cooperated	… defected	P-value (χ^2 test)
1	Costly, equal	15.1%	7.0%	0.038
2	Costly, unequal	3.0%	7.8%	0.072
3	Selfish, equal	85.3%	89.0%	0.406
4	Selfish, unequal	52.6%	59.4%	0.259

Table 3: Unkind decisions by subject's partner's action in the Prisoner's Dilemma

Now, we examine our main question: if their PD partners defected, did subjects react differently to the two groups? To do this, we compare subjects' behaviour to the Partner group and the Other group. Table 4 shows the percentage of unkind decisions, disaggregated into decisions against (members of) the Partner and Other groups, and by whether the subject's Prisoner's Dilemma partner played Cooperate or Defect. Figure 1 shows the same information graphically.

		PD partner cooperated		PD partner defected	
Choice sets	Unkind decision is...	Partner group	Other group	Partner group	Other group
1	Costly, equal	12.1%	18.1%	9.4%	4.7%
2	Costly, unequal	4.3%	1.7%	7.8%	7.8%
3	Selfish, equal	87.1%	83.6%	89.1%	89.1%
4	Selfish, unequal	53.4%	51.7%	65.6%	53.1%

Table 4: Unkind decisions by PD partner's action and target group

To analyse this data, we look at subjects' choices in each *pair* of decisions, split up by whether their PD partner defected or cooperated. In this way we can use the within-subject aspect of our design to see if subjects took decisions differently depending on whether they

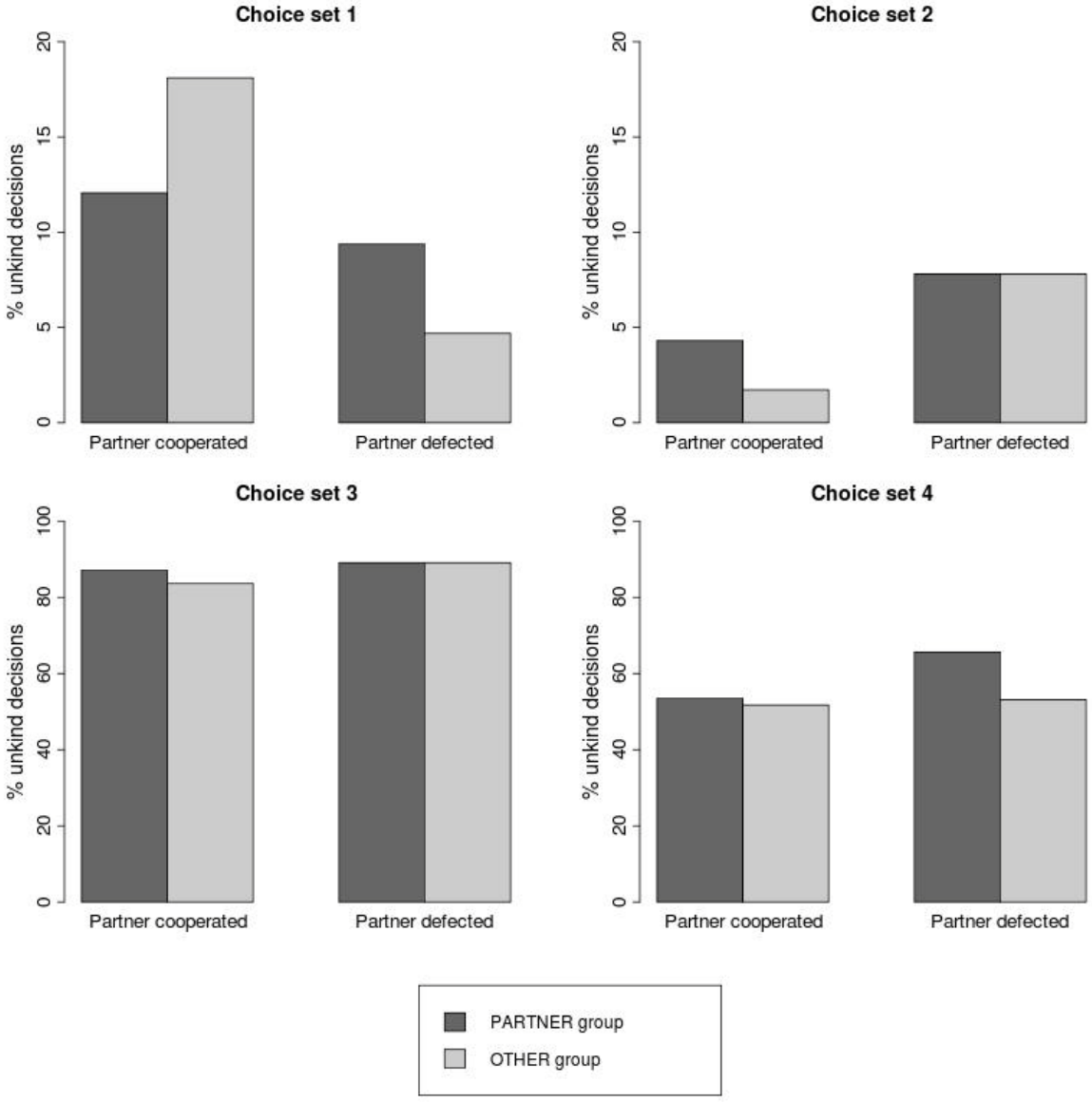

Figure 1: Unkind decisions by choice set and PD partner's action

were facing the Partner group or the Other group. Table 5 gives the raw data. For each pair of decisions, we report P-values for a binomial test of whether subjects were significantly more unkind to the Partner than to the Other group. The test uses "discordant pairs" – subjects who were unkind to the Partner group member but not the Other group member, or subjects who were unkind to the Other group member but not the Partner group member. Under the null hypothesis, the proportion of all discordant pairs in either one of these two categories is binomially distributed with parameter 0.5. Thus, when there are many more subjects being

unkind to the Partner group only, rather than vice versa, the null will be rejected. (However, we report two-tailed significance tests for consistency with other tests in the paper.)

	Choice sets 1/6: PD partner...		Choice set 2: PD partner...		Choice set 3: PD partner...		Choice set 4: PD partner...	
Choice for Partner group/Other group	Cooperated	Defected	Cooperated	Defected	Cooperated	Defected	Cooperated	Defected
Kind/kind	94 (81.0%)	56 (87.5%)	111 (95.7%)	56 (87.5%)	9 (7.8%)	5 (7.8%)	47 (40.5%)	19 (29.7%)
Kind/unkind	**8 (6.9%)**	**2 (3.1%)**	**0 (0.0%)**	**3 (4.7%)**	**6 (5.2%)**	**2 (3.1%)**	**7 (6.0%)**	**3 (4.7%)**
Unkind/kind	**1 (0.9%)**	**5 (7.8%)**	**3 (2.6%)**	**3 (4.7%)**	**10 (8.6%)**	**2 (3.1%)**	**9 (7.8%)**	**11 (17.2%)**
Unkind/unkind	13 (11.2%)	1 (1.6%)	2 (1.7%)	2 (3.1%)	91 (78.4%)	55 (85.9%)	53 (45.7%)	31 (48.4%)
P-value (binomial test)	0.039*	0.45	0.250	1	0.454	1	0.804	0.0574+

Column percentages in parentheses. P-value is two-tailed.
*** p < 0.001; ** p < 0.01; * p < 0.05; + p < 0.10.

Table 5: Pairs of decisions, by PD partner's choice

There are two statistically significant results.

Result 1: When the unkind decision was costly and egalitarian, subjects were significantly kinder to the Partner group if their PD partner had cooperated, but were not significantly kinder if their partner had defected.

First, in choice set 1, where the unkind decision was costly and egalitarian, if the PD partner cooperated, subjects were significantly less likely to choose the unkind decision for the Partner group than for the Other group (p=0.039). There was no significant difference if the PD partner defected. Thus, we must qualify the previous claim that cooperation by the PD partner increased egalitarianism. It appears to have only done so for the Other group. One interpretation is that there were countervailing motives: the PD partner's defection decreased

the willingness to be egalitarian, but also increased the willingness to harm those in the partner's group.

Result 2: When unkindness was selfish and inegalitarian, defection by a subject's PD partner increased the subject's unkindness to the Partner group, relative to the Other group.

Second, in choice set 4 where the unkind decision was selfish and unequal, if the PD partner defected, subjects were more likely to choose the unkind decision for the Partner group than for the Other group. The increase in selfishness towards the Partner group after partner's defection is quite large (from 53.4% to 65.6%), although significance is weak (p=0.057).

In a final non-parametric test, we sum “discordant pairs” over all five choice sets. After the PD partner defected, a total of 31 decision pairs discriminated against the Partner group, while 17 discriminated against the Other group (binomial test, p=0.059). After the PD partner cooperated, 41 decision pairs discriminated against the Partner group, while 33 discriminated against the Other group (binomial test, p=0.416).

Regression models

We examine choices further by modelling each pair of choice sets as a bivariate probit, allowing subjects' decisions to be correlated within the pair. The independent variable, as before, is whether the PD partner defected. We report results only for sets 1 and 4: other choice sets gave null results as before. The dependent variables take the value of 1 if the subject chose the unkind decision, and 0 otherwise, for the Partner group and for the Other group respectively.

	Choice set 1	Choice set 4

Partner: intercept	-1.203 (0.152) ***	-0.088 (0.117)
Partner: PD partner defected	-0.080 (0.261)	0301 (0.198)
Other: intercept	-0.919 (0.136) ***	0.043 (0.116)
Other: PD partner defected	-0.644 (0.281) *	0.041 (0.195)
Correlation (ρ)	0.880	0.878
N	180	180

*** p < 0.001; ** p < 0.01; * p < 0.05; + p < 0.10.

Table 6: Binomial probit regressions

The results for choice set 1 confirm that the PD partner's defection decreased the probability of taking the unkind decision only for the Other group. For choice set 4, the effect of the PD partner's decision just misses 10% significance for the Partner group and is insignificant for the Other group. The significance and signs of parameters can be misleading in non-linear models, so Table 7 shows the effect of PD partner's defection in terms of probabilities.[8] In both cases, defection increases the probability of being unkind to the Partner group *only*, and decreases the probability of being unkind to the Other group *only*. This is what one would expect if subjects were group-reciprocating. Looking at the effect on unkindness to each group in general (i.e., including the possibility that the subject was unkind to both groups), defection makes subjects kinder in choice 1, and unkinder in choice 4, but the size of the effects varies between the groups by roughly an order of magnitude: subjects get much kinder to the Other group, and much unkinder to the Partner group, in the respective decisions. Lastly we report the p-value of the null hypothesis that the first two rows are equal, i.e. that subjects' behaviour to each group is affected in the same way by PD partner defection. (This is also the p-value of the equivalent hypothesis that the last two rows are equal.) The null is strongly rejected for choice set 1, but cannot be rejected for set 4.

8 The raw probabilities are shown in Table 9 in the Appendix.

Difference in % probability when PD partner defected	Choice set 1	Choice set 4
1. Unkind to Partner group only	3.37 (2.67)	6.13 (4.18)
2. Unkind to Other group only	-7.15 (2.53)**	-3.87 (2.24)+
3. Unkind to Partner group (any decision for Other group)	-1.47 (4.74)	11.63 (7.53)
4. Unkind to Other group (any decision for Partner group)	-11.99 (4.58)**	1.63 (7.77)
Prob(1=2)	0.015	0.111

Standard errors in parentheses (calculated by the delta method).
*** p < 0.001; ** p < 0.01; * p < 0.05; + p < 0.10.

Table 7: Change in probabilities of unkind decisions, from bivariate probits

Interpretation

Due to the lack of high statistical significance, the results obtained above need to be handled with care. Nevertheless, we can form some tentative hypotheses.

First, group reciprocity did not affect all decisions equally. Effects were only significant or large when decisions pitted self-interest against inequality concerns. When subjects were asked to make a straight allocation between Partner and Other group members, with their own payoffs not involved, no group discrimination was observed. This suggests to us that reciprocity operates via changes in self-other inequality concerns, rather than via changes in altruism or spitefulness towards different groups. Indeed, if group spite had been a prominent motivation, it ought to have been so in choice set 3 as well – which was not observed.

In choice set 1, self-interest conflicted with a dislike of other-favouring inequality: choosing the best outcome for yourself meant giving the other subject 3 Euros more than you. In general, subjects were more likely to choose this outcome after their partners defected in the

PD; but this increase in generosity (or self-interest) did not occur when making decisions against the Partner group.

It seems counter-intuitive that decisions concerning the Other group should be changed more by PD partner's defection than decisions concerning the Partner group. There are two possibilities. (1) We may have observed two countervailing effects. Subjects may have feared leaving the experiment with low earnings, or experiencing defection lessened their attachment to fairness norms, and this made them more self-interested in general; on the other hand, some subjects also wished to punish the partner group. (2) Subjects might have a specific “desire to discriminate” which could result in them becoming simultaneously nastier to their PD partner's group, and kinder to the other group. Although the data cannot differentiate these explanations, we favour the first.

In choice set 4, self-interest conflicted with a dislike of self-favouring inequality: choosing the best outcome for yourself meant giving the other subject much less than yourself. Here, more straightforwardly, experiencing defection made subjects more likely to be unkind to the PD partner's group and less likely to be unkind to the Other group. Significance levels for these choices are not so high, however.

In many circumstances, discriminatory behaviour may be driven by beliefs rather than by preferences. For example, ethnic discrimination in hiring may come from “tastes” (Becker 1957), or from statistical or screening discrimination (Arrow 1972). Similarly, Yamagishi et al. (2000) suggest that ingroup favouritism happens because people believe that other group members will reciprocate their nice behaviour. However, in our experiment, changes in beliefs about future behaviour are ruled out: players simply make allocation decisions and

know that other players cannot react to them.[9] Thus, the behaviour we observe must come from preferences, not beliefs. However, we caution against too narrow an interpretation of preferences. Group reciprocity may be due to the activation of context-specific norms, rather than to "preferences" thought of as unchanging characteristics belonging to individuals.

Conclusion

Most economic models assume that humans are rational and materially self-interested. In many areas of economic life, evidence from laboratory and the field has cast doubt on these assumptions. One of the most important differences between real humans and those in economic models, we suspect, is that real humans act in groups, and treat other groups as actors. As a result they may assign blame and praise to groups as well as individuals – with important consequences for human society and politics. Often in history, entire groups have been scapegoated and blamed for their members' real or imagined behaviour. It is important to

9 We cannot rule out that this behaviour is caused by "counterfactual reciprocity", where subjects take action because of what they expect others *would have done*, e.g. if their positions had been reversed. In fact, it is hard to rule this out completely for any kind of decision. However, we can rule out expectations about actual future behaviour.

understand how this happens. Using arbitrary, lab-created groups, in an experiment using real money payoffs, we showed that subjects responded to behaviour by one member of a group by treating other members of that group differently. There was evidence for reciprocity both when harming the other group was materially beneficial (choice set 4) and when it was costly (choice set 1). Thus, group reciprocity seems to be part of human psychology.

Not all our subjects showed group reciprocity, and it was not visible in all decisions. To investigate the causes and correlates of group reciprocity, our experimental paradigm needs further refinement. In particular, continuous rather than discrete decisions might allow for more accurate measurement of individuals' levels of reciprocity. Complementary experiments could also be done using homegrown identities such as ethnicity. And for external validity, theories derived from the lab will need to be tested against behaviour in real world group conflicts. Finally, we have still to investigate third-party group reciprocity: when people retaliate against a group for harm done not to themselves, but to a fellow group member of their own.

We hope that demonstrating group reciprocity in a reproducible laboratory experiment will open the door to deeper exploration of how it works. We can draw an analogy with the “minimal group paradigm” in psychology. After Shaerif's (1966) initial research in the 1950s on in-group prejudice – the famous Robber's Cave experiment – researchers developed a canonical experiment (“paradigm”) to analyse in-group prejudice. By making the behaviour of interest reproducible, the minimal group paradigm catalysed much productive research into prejudice. Many kinds of group reciprocal behaviour – racial prejudice, discrimination and conflict – belong to the extremes of human conduct. However, it is not necessarily true that extreme situations have their own special psychology. Group reciprocity may happen not only in Yugoslavia or Rwanda, but also in ordinary life, between football supporters or the

employees of rival firms. If so, then it ought to be replicable in the laboratory. Therefore, we hope that other researchers will take up and refine our experimental design.

Bibliography

Abbink, K., and B. Herrmann. 2009. Pointless Vendettas. Manuscript.

Archer, John (2004), "Sex Differences in Aggression in Real-World Settings: A Meta-Analytic Review", *Review of General Psychology* 8(4): 291-322.

Arrow, Kenneth J. (1972) "Some mathematical models of race discrimination in the labor market," in: Pascal, Anthony H. (ed.), *Racial discrimination in economic life*, Lexington: Lexington Books: 187-203.

Bauerlein, M. 2001. *Negrophobia: A race riot in Atlanta, 1906*. Encounter Books.

Becker, Gary S. (1971), *The economics of discrimination*, University of Chicago Press.

Bornstein, Gary (1992), "The Free Rider Problem in Intergroup Conflicts over Step-Level and Continuous Public Goods", *Journal of Personality and Social Psychology* 62: 597-606.

Bornstein, Gary (2003), "Intergroup Conflict: Individual, Group, and Collective Interests", *Personality and Social Psychology Review* 7(2): 129-145.

Brewer, Marilynn B. (1979), "Ingroup Bias in the Minimal Intergroup Situation: A Cognitive-Motivational Analysis", *Psychological Bulletin* 86, 307-324

Brewer, Marilynn B. 1999. The Psychology of Prejudice: Ingroup Love or Outgroup Hate? *Journal of Social Issues* 55, no. 3 (Fall99): 429. doi:Article.

Chen, Yan and Sherry Xin Li (2009), "Group Identity and Social Preferences", *American Economic Review* 99(1): 431-457.

De Cremer, David and van Vugt, Mark (1999), "Social Identification Effects in Social Dilemmas: A Transformation of Motives", *European Journal of Social Psychology* 29: 871-893.

Gaertner, Lowell and John Schopler (1998), "Perceived Ingroup Entitativity and Intergroup Bias: An Interconnection of Self and Others", *European Journal of Social Psychology* 28: 963-980.

Guala, Francesco, Luigi Mittone and Matteo Ploner (2009), "Group Membership, Team Preferences, and Expectations", University of Trento, *CEEL Working Papers* 0906.

Hargreaves Heap, Shaun and Daniel J. Zizzo (2009), "The value of groups", *American Economic Review* 99(1): 295-323

Havely, Nir, Gary Bornstein and Lilach Sagiv (2008), "`In-Group Love´ and `Out-Group Hate´ as Motives for Individual Participation in Intergroup Conflict", *Psychological Science* 19(4): 405-411.

Horowitz, D. L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Horowitz, D. L. 2001. *The deadly ethnic riot*. University of California Press.

Kocher, M. A., T. B. Pepinsky, and S. N. Kalyvas. 2008. Into the Arms of the Rebels? Aerial Bombardment, Indiscriminate Violence, and Territorial Control in the Vietnam War. Paper prepared for the ISA's 49th ANNUAL CONVENTION, Hilton San Francisco, SAN FRANCISCO, CA, USA.

Lickel, B., N. Miller, D. M. Stenstrom, T. F. Denson, and T. Schmader. 2006. Vicarious retribution: The role of collective blame in intergroup aggression. *Personality and Social Psychology Review* 10, no. 4: 372.

Mamdani, M. 2001. *When victims become killers: colonialism, nativism, and the genocide in Rwanda*. Princeton University Press Princeton.

Moreno, A. T. 2008. Group Fairness and Game Theory. *School of Economics Working Papers*.

Mummendey, A., and S. Otten. 1998. Positive–negative asymmetry in social discrimination. *European review of social psychology* 9, no. 1: 107-143.

Reuben, Ernesto and Frans van Winden (2008), "Social Ties and Coordination on Negative Reciprocity: The Role of Affect", *Journal of Public Economics* 92: 34-53.

Ritchie, D. 2001. Oldham independent review: one Oldham, one future. *Manchester: Government Office for the North West*.

Scheff, Thomas (2003), "Male Emotions/Relationships and Violence: A Case Study", *Human Relations* 56(6): 727-749.

Shayo, M., and A. Zussman. 2010. Judicial Ingroup Bias in the Shadow of Terrorism.

Sherif, M., O. J. Harvey, B. J. White, W. R. Hood, and C. W. Sherif. 1961. *The Robbers Cave experiment: Intergroup conflict and cooperation*. Wesleyan University Press Scranton, Pa.: Distributed by Harper & Row, Middletown, Conn.

Sorkin, Andrew Ross. 2009. *Too Big to Fail: The Inside Story of How Wall Street and Washington Fought to Save the Financial System---and Themselves*. Viking Adult, October 20.

Staub, Ervin (2000), "Genocide and Mass Killing: Origins, Prevention, Healing and Reconciliation", *Political Psychology* 21(2): 367-382.

Stenstrom, Douglas M., Brian Lickel, Thomas F. Denson and Norman Miller (2008), "The Roles of Ingroup Identification and Outgroup Entitativity in Intergroup Retribution", *Personality and Social Psychology Bulletin* 34(11):1570-1582.

Yamagishi, Toshio and Toko Kiyonari (2000), "The Group as the Container of Generalized Reciprocity", *Social Psychology Quarterly* 63(2): 116-132.

Yamagishi, Toshio and Nobuhiro Mifune (2008), "Does Shared Group Membership Promote Altruism?", *Rationality and Society* 20(1): 5-30.

Yamagishi, Toshio and Nobuhiro Mifune (2009),"Social Exchange and Solidarity: In-Group Love or Out-Group Hate?", *Evolution and Human Behavior* 30: 229-237.

Appendix: tables and figures

N	180
Demographics	
Male	142 (78.89%)
Studying:	
… Law	8 (4.44%)
… Social sciences	56 (31.11%)
… Economics	27 (15.00%)
… Natural sciences	38 (21.11%)
… Other	38 (21.11%)
… Not a student	13 (7.22%)
Age	Min 19, max 67, mean 24.6, median 24

Prisoner's Dilemma choices (self, other)

CC	72 (40.00%)	DC	44 (24.44%)
CD	44 (24.44%)	DD	20 (11.11%)

Allocations (those choosing option 2)

Choice set 1P	20 (11.11%)	Choice set 1O	24 (13.33%)
Choice set 2P	10 (5.56%)	Choice set 2O	7 (3.89%)
Choice set 3P	158 (87.78%)	Choice set 3O	154 (85.56%)

Choice set 4P	104 (57.78%)	Choice set 4O	94 (52.22%)
Choice set 5P	71 (39.44%)	Choice set 5O	80 (44.44%)

Group identity Likert scales (1=not at all, 7=very much)

It is fun to be part of the COLOUR group	≤ 2: 49 (27.22%); ≥ 6: 24 (13.33%); mean 3.66
It feels good to be part of the COLOUR group	≤ 2: 66 (36.67%); ≥ 6: 20 (11.11%); mean 3.34
I am happy to be in the COLOUR group	≤ 2: 51 (28.33%); ≥ 6: 22 (12.22%); mean 3.56
I think members of the COLOUR group have a lot to be proud of	≤ 2: 53 (29.44%); ≥ 6: 18 (10.00%); mean 3.57
I feel solidarity with members of the COLOUR group	≤ 2: 39 (21.67%); ≥ 6: 33 (18.33%); mean 3.99
I feel connected to the COLOUR group	≤ 2: 49 (27.22%); ≥ 6: 34 (18.89%); mean 3.83

Table 8: Descriptive statistics

	Choice sets 1/4: PD partner...		Choice sets 6/9: PD partner...	
Choice for Partner group/Other group	Cooperated	Defected	Cooperated	Defected
Kind/kind	80.2%	88.8%	39.4%	31.7%
Kind/unkind	8.4%	1.2%	7.1%	3.2%
Unkind/kind	1.9%	5.3%	8.8%	15.0%
Unkind/unkind	9.5%	4.7%	44.7%	50.2%

Table 9: Predicted probabilities from binomial probits