

5-2017

Gender, Punishment, and Cooperation: Men Hurt Others to Advance Their Interests

Terence C. Burnham
Chapman University, burnham@chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/esi_working_papers



Part of the [Econometrics Commons](#), [Economic Theory Commons](#), and the [Other Economics Commons](#)

Recommended Citation

Burnham, T. C. (2017). Gender, punishment, and cooperation: Men hurt others to advance their interests. ESI Working Papers 17-11. Retrieved from http://digitalcommons.chapman.edu/esi_working_papers/221

This Article is brought to you for free and open access by the Economic Science Institute at Chapman University Digital Commons. It has been accepted for inclusion in ESI Working Papers by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

Gender, Punishment, and Cooperation: Men Hurt Others to Advance Their Interests

Comments

Working Paper 17-11

Gender, Punishment, and Cooperation:
Men hurt others to advance their interests

May 2017

© Terence C. Burnham

Associate Professor
Chapman University
One University Avenue, Orange, California, USA
burnham@chapman.edu
714/744-7847, 714/532-6081 (fax)

Abstract

A laboratory experiment that reports on gender, cooperation, and punishment in two repeated public goods game using high-powered punishment. In a repeated public goods game with punishment, no statistically significant differences between men and women are reported. In a modified game that adds an explicit payoff for relative performance, men punish more than women, men obtain higher rank, and punishment by males decreases payoffs for both men and for women. These results contribute to the debate about the origins and maintenance of cooperation.

Keywords

Cooperation, Reciprocity, Punishment, Public-Goods, Altruism, Gender

JEL classification

A13; C72; C91

1 Introduction

Cooperation is a central issue for the social sciences, the study of animal behavior, and myriad other fields. There is an active, and unresolved, debate regarding the evolutionary origin of punishment, and its role in fostering cooperation (Trivers 1971, Yamagishi 1986, Clutton-Brock and Parker 1995, Fehr and Gächter 2000, Fehr and Gächter 2002, Fowler 2005, Sigmund 2007).

Punishment has been the focus of considerable research because it can increase cooperation in some settings (Yamagishi 1986, Fehr and Gächter 2002). The nature of punishment varies between societies; in some societies, which may be unusual (Henrich, Heine et al. 2010), punishment is preferentially imposed upon relatively low cooperators. Subjects drawn from other populations, however, exhibit 'anti-social' punishment that is directed toward high cooperators (Shinada, Yamagishi et al. 2004, Herrmann, Thöni et al. 2008). Punishment can increase efficiency (in situations where the cost of punishment is lower than the benefits created by increased cooperation), although not in all settings (Rand and Nowak 2011, Burnham 2014).

Cooperation, punishment, and reciprocity are also important to a wide-range of non-human species including apes (Brosnan, Schiff et al. 2005, Silk, Brosnan et al. 2005, Jensen, Hare et al. 2006, Jensen, Call et al. 2007, Jensen, Call et al. 2007, Vonk, Brosnan et al. 2008, Brosnan, Silk et al. 2009, Van Leeuwen, Zimmermann et al. 2011, Riedl, Jensen et al. 2012, Proctor, Williamson et al. 2013), monkeys (Brosnan and de Waal 2003, Cronin, Schroeder et al. 2010, Talbot, Freeman et al. 2011), and rodents (Rutte and Taborsky 2007). Cooperation and reciprocity may impact both hierarchy and health outcomes (Sapolsky 2005, Yee, Cavigelli et al. 2008).

Beginning in 1964, various types of cooperative behavior that had been considered puzzles have been reconciled with evolutionary theory via kin selection (Hamilton 1964), reciprocal altruism (Trivers 1971), costly signaling (Zahavi 1975), group selection (Wilson 1975), and indirect reciprocity (Nowak and Sigmund 1998).

These theories reveal costly cooperative behavior as a means to increase the inclusive fitness of the cooperator, thus resolving any perceived paradox. Under reciprocal altruism, indirect reciprocity, and costly signaling, the individual organism receives expected future benefits that more than compensate for the cost of cooperation; in kin selection and group selection, the individual organism need not be repaid, but its genes increase in relative frequency via the impact on other organisms.

While there has been enormous progress in understanding cooperation, one area remains controversial: costly cooperation in situations where the behavior cannot redound to the benefit of the cooperators' genes. In anonymous and final interactions with anonymous strangers, costly cooperation cannot redound to the benefit of the cooperators' genes via any of the five theories listed above (Burnham 2014).

1. If the players are not kin, then costly acts cannot redound to the benefit of genes in kin via kin selection.
2. If the players are not part of evolutionarily relevant groups, then costly acts cannot redound to the benefit of genes via group selection.
3. If there are no future interactions, then costly acts cannot be repaid via reciprocal altruism.
4. If no reputation can be formed, then costly acts cannot be repaid via indirect reciprocity.
5. If behavior is anonymous then costly acts cannot be repaid via signaling.

Thus, the remaining 'puzzle of cooperation' is the human propensity to cooperate even in final and anonymous interactions played among people who are neither kin nor members of evolutionarily relevant groups.

There are two main explanations for this 'puzzling' cooperation. The first is 'strong reciprocity', a form of group selection (Gintis 2000, Fehr, Fischbacher et al. 2002). Note that strong reciprocity is not standard group selection as the individuals in laboratory experiments are not part of evolutionarily relevant groups (Burnham and Johnson 2005). Thus, strong reciprocity is a new aspect of human nature where the costly cooperative behavior is benefits anonymous, transitory strangers in laboratories (Burnham 2014).

Strong reciprocity is claimed to be a new, sixth explanation for costly cooperative behavior. If strong reciprocity exists it means that economics, biology, and other fields would have to be re-written from their axioms.

A number of scholars argue against the existence of 'strong reciprocity' (Price, Cosmides et al. 2002, Johnson, Stopka et al. 2003, Trivers 2004, Burnham and Johnson 2005, Binmore 2006, Dreber, Rand et al. 2008, West, Mouden et al. 2011, Burton-Chellew and West 2013, Kurzban and DeScioli 2013, Krasnow, Delton et al. 2015).

Although these critiques of strong reciprocity vary in their explanations, three common themes emerge. First, the documented costly behavior is real, at least in the laboratory. The experimental facts are not in dispute. Second, group selection is not the preferred explanation. Third, a mismatch between laboratory and non-laboratory settings may explain the inability to identify benefits to the costly cooperative behaviors anonymous and final interactions within transient groups of strangers. Behaviors that are, or were in ancestral settings, paid back outside the laboratory may appear puzzling inside the laboratory.

This study examines gender and punishment to increase our understanding of cooperation, and to help resolve the competing views on so-called 'strong reciprocity.'

The first motivation for this work is practical. To the extent that women and men behave differently, institutions designed to foster cooperation may need to be different depending on the gender composition of the group. For example, primary schools that only educate girls or boys might require different institutions to obtain the same goal. Beyond schools, there are a variety of settings with extreme gender distributions including the military, prisons, and some careers (e.g., pediatrics). When it comes to gender and cooperation, one approach might not work in all settings.

The second motivation for this work is theoretical. Gender differences may inform the debate regarding the existence of strong reciprocity. According to strong reciprocity, punishment is altruistic. Thus, if we look at the characteristics of punishers, those characteristics should align with theories of altruism.

In contrast, the anti-strong reciprocity prediction is that punishment in laboratory settings is administered by people who would punish in non-laboratory settings, either modern or ancestral.

In support of this anti-strong reciprocity view, across many species, punishment is used by individuals for personal dominance, to discourage cheating, to discipline conspecifics, and to maintain cooperation (Clutton-Brock and Parker 1995, Raihani, Thornton et al. 2012). The punishers incur costs in the act of punishment, and reap future benefits that exceed the costs (in expectation). Punishing behavior is linked to dominance seeking behavior in a wide-variety of primates including humans (Mazur and Booth 1998, Mazur 2015). There is nothing altruistic about punishment and dominance-seeking in this view.

What is known about gender and cooperation? Existing studies reveal no clear relationship between gender and certain cooperative behaviors. In public goods games (without punishment), no 'systematic' difference is reported between men and women (Eckel and Grossman 2008). Moreover, a review article reports no consistent relationship between cooperation and gender in prisoners' dilemma and social dilemma games (Croson and Gneezy 2009).

Notwithstanding the published work, in contexts relevant to the debate over strong reciprocity, men may punish more than women for two reasons.

First, punishment may be viewed as similar to physical conflict (and outside the laboratory even monetary punishment may be counter-punished by physical violence). Men are argued to favor physical punishment of unfair behavior (Singer, Seymour et al. 2006), and punishment is reported to vary between patriarchal and matriarchal societies (Asiedu and Ibanez 2014).

Gender effects on rejection behavior in the ultimatum game are reported to vary based on the design in a way that is consistent with this hypothesis (Croson and Gneezy 2009). In a face-to-face ultimatum game, women are reported to reject less than men (Eckel and Grossman 2001). However, in a different study without direct visual contact, women are reported to reject more than men (Solnick 2001). Presumably, there was no possibility of physical violence within either of these experiments, but face-to-face interactions might include more cues to possible violence.

Second, women may care less about rank than men. If so, punishment may be a tool to advance in relative rank (Burnham 2015). When given a choice, women choose a competitive tournament setting less than men (Niederle and Vesterlund 2007), although in some cultures women behave more competitively than men (Gneezy, Leonard et al. 2008). More broadly, men are reported to be less cooperative (Molina, Giménez-Nadal et al. 2013), and less generous (Selten and Ockenfels 1998). Both men and women reveal a preference for rank (Tran and Zeckhauser 2012), and in a field experiment, women are less 'heedful' of rank than men (Barankay 2011).

With these motivations, this experiment examines gender differences in punishment in two repeated public goods games – one with explicit monetary incentives to obtaining high rank, and a second with such monetary incentives.

2 Methods

A total of 96 undergraduate students from Chapman University voluntarily participated in the experiments. Four experimental sessions with 24 subjects each were conducted. Each session had equal numbers of men and women for a total of 48 women and 48 men.

All the public goods games in this experiment use high-powered punishment that inflicts 50 units of monetary punishment for each 1 unit of cost to the punishment (50:1 hereafter). 50:1 may be more representative of non-laboratory payoffs both in industrialized and ancestral settings (Wiessner 2005, Burnham 2015).

Subjects in each session played a repeated public goods game with punishment and a second public goods game identical to the first except for the addition of explicit payoffs for relative performance (Burnham 2015). The order of the treatments was counterbalanced between sessions so that half the subjects played the game with relative performance payoffs first.

In each repeated public goods game, the 24 subjects were allocated to six groups of four subjects. Groupings were structured in the manner of the ‘perfect stranger’ treatment of Fehr and Gächter (2002). The allocation of subjects to the groups ensured that, within a given treatment, no subject ever met the same person more than once. In each round, subjects were identified with a transient identifier to ensure no reputations could be formed. At the end of each period, subjects were informed about their own decisions, the decisions of the other group members, and their payoff in current period.

Subjects were given written instructions that explained the structure of the game, the composition of groups in each period, the inability to form reputations because of anonymity. After the instructions, and before the experiment, subjects were given a test of knowledge on several hypothetical examples. In order to participate, each subject was required to get all the payoff examples correct. All experimental decisions were made on a computer screen using z-tree software.

The subjects sat in four rows with six individuals per row. All decisions were made via computers, and each subject had her or his own computer. 3-sided opaque screens separated each computer and subject. Subjects were instructed not to look at anyone else’s screen and not to speak to each other.

Subjects each received \$7 in advance for participation, and were paid according to behavior in the game. In each round, each player was given 20 Experimental Currency Units (ECU) to allocate between a public and a private account. ECUs in the private account remained with the player, while those allocated to the group account were multiplied by 1.6 and divided equally among 4 players in a group. ECUs were converted to cash at a rate of 100 tokens to \$1. Subjects were paid in cash and privately at the end of the session.

The punishment phase came after each round of the public good. Group members were identified by a transient number, and by their contribution to the public good. Each player could allocate up to 10 units of punishment to each of the 3 other group members. A unit of punishment cost the punisher 1 ECU and reduced the punished player's payoff by 50 ECU.

As noted above, each subject played two repeated public goods games with punishment. One of these two games including monetary payoffs based on relative performance that pay high-ranking subjects a bonus, and impose a penalty on low-ranking subjects. The total, and average, of the rank-based, extra payments was zero (Burnham 2015).

The specific relative performance payoffs were determined as follows. At the end of the 6-period repeated public goods game with punishment, the 24 subjects in the session were ranked from 1st to 24th. 1st place was given a bonus of 550 ECUs. 2nd place earned 500 ECU's. Payoffs decreased by 50 per place except that rank of 12 and rank of 13 each earned 0 ECUs.

The 50:1 punishment technology allows for large negative payouts. The maximum punishment possible in a round is 30 points (10 units of punishment from each of the 3 other players in the group), which would result in a loss of 1,500 ECU. Players with negative balances would no longer suffer the costs of punishment, and might behave differently. Accordingly, two changes were made. First, endowments were large relative to the public goods game payoffs; endowments in three sessions were 2000 ECU and in one session the endowment was 1500 ECU. Second, the game would be stopped if any player's cumulative balance turned negative.

Subjects were recruited by email via Chapman University's Economics Science Institute (ESI) email list. Subjects were restricted to not having participated in previous ESI public goods experiments. All sessions were conducted in May 2012, beginning at the same time (4pm) on a Tuesday, Wednesday, or Thursday.

No deception was used in the experiment.

3 Results

3.1. Contributions

Both women and men contribute almost maximally to the public good, and there is no statistically significant difference between women and men in either treatment (table 1, figure 1).

In the 'standard' public goods game women contribute an average of 17.03/20 per round versus 17.98/20 for men. In the public goods game with rank-based payoffs, women contribute an average of 18.0/20 versus 18.64/20 for men. These differences are not statistically significant.

	PG	PG w/rank
Women	17.03	18.00
Men	17.98	18.64
std-err, women	0.42	0.42
Std-err, men	0.39	0.47

p-value (2-tailed)	0.10	0.32
--------------------	------	------

Table 1: Contributions to the public good by treatment and gender

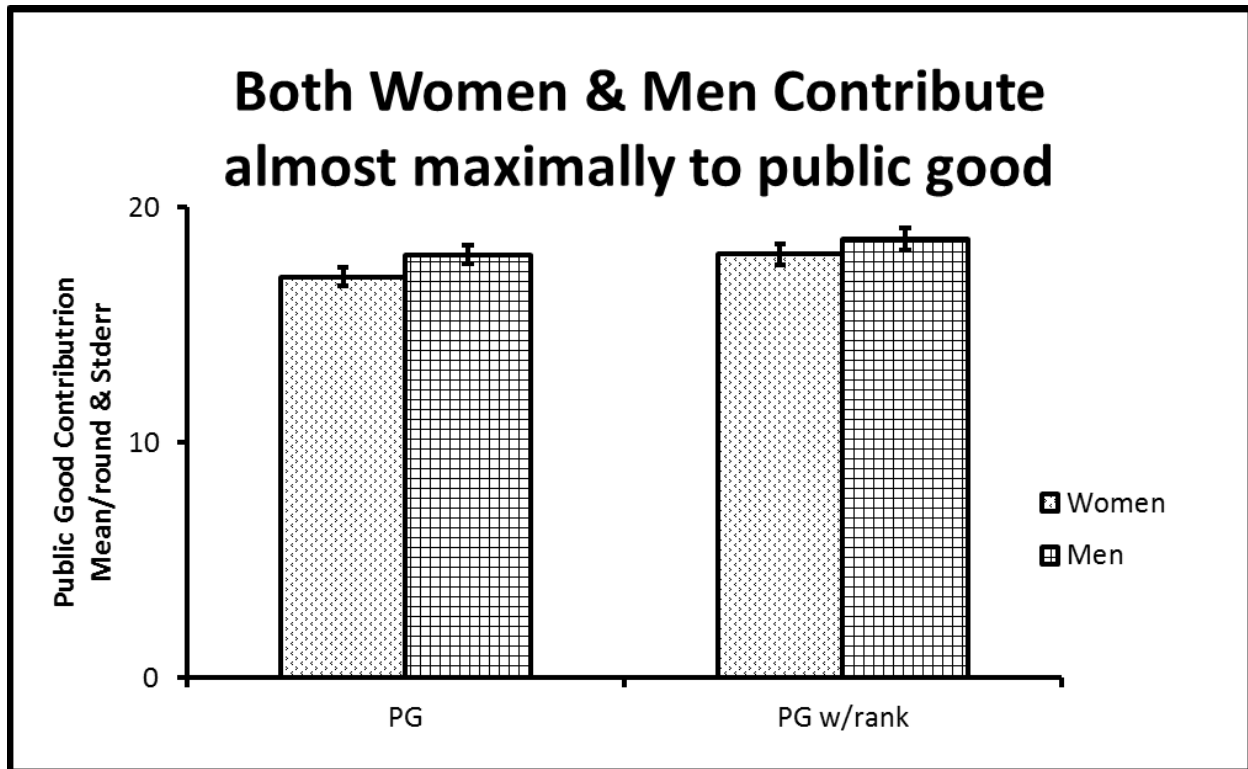


Figure 1: Contributions to the public good by treatment and gender

3.2 Punishment

There is no statistically significant difference in punishment by men and women in the 'standard' public goods game. Women administer 1.76 units punishment per round out of a maximum possible punishment of 30 units (10 units administered to each of the three other players in the group) while men administered an average of 2.46 punishment units.

In contrast, in the presence of rank-based payoffs, men administer more than twice as much punishment. Women administer an average of 4.34 units punishment per round out of a maximum possible punishment of 30 units (10 units administered to each of the three other players in the group) while men administer an average of 8.74 punishment units per round. This difference is significant at the $p < 0.01$ level.

	PG	*PG w/rank
Women	1.76	4.34

Men	2.46	8.74
std-err, women	0.42	0.68
Std-err, men	0.50	1.12
p-value (2-tailed)	0.29	0.00

Table 2: Punishment administered by treatment and gender

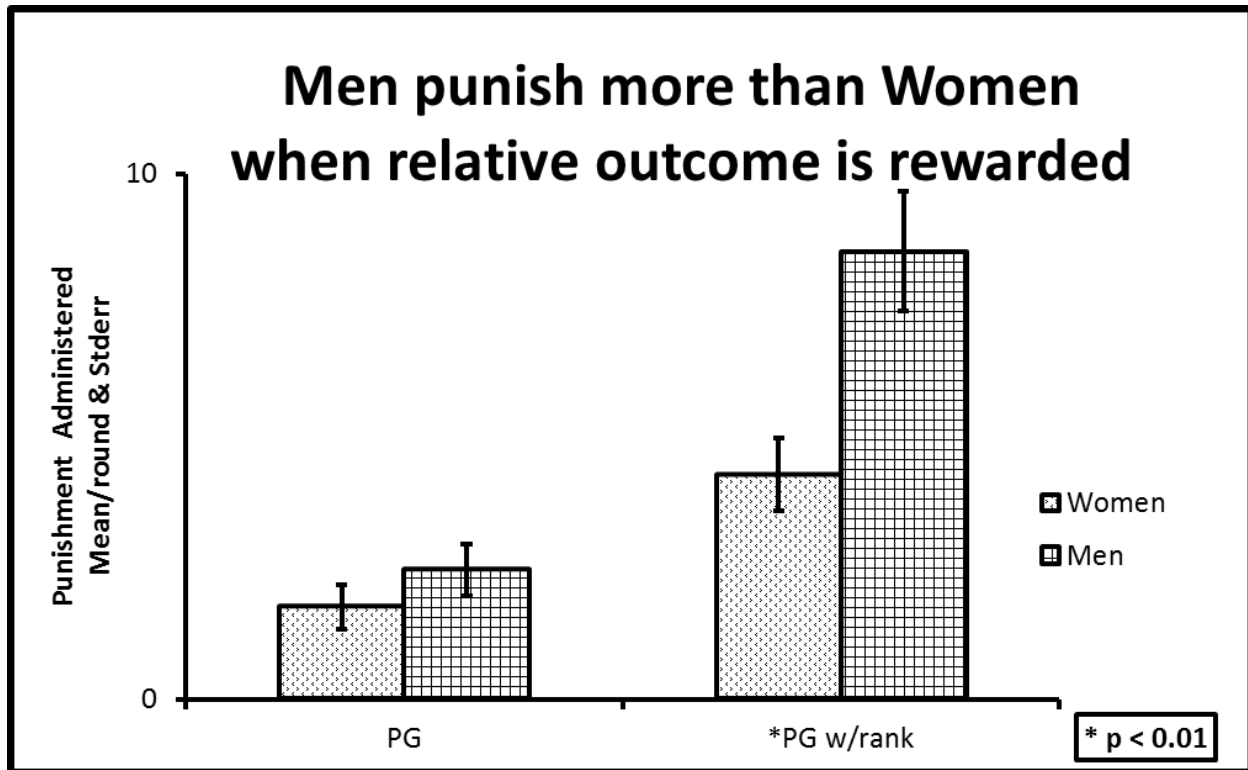


Figure 2: Punishment administered by treatment and gender

3.3 Rank

At the end of each public goods game, the 24 subjects are ranked from 23 to 0 based on earning in that PG game. A ranking of 23 is the highest possible and means that the 23 other subjects ranked below this individual. This ranking system has a larger number equating to higher rank.

There is no statistically significant difference in rank for men and women in the standard public goods game. Women obtain an average rank of 11.04 while men obtain a statistically indistinguishable 11.96.

In the presence of rank-based payoffs, however, men obtain significantly higher rank than women. Women obtain an average rank of 10.04, almost 3 full places below the 12.96 obtained by men. This gender difference is significant that the $p < 0.05$ level.

This difference in rank of almost 3 slots between men and women is large. The maximum possible difference with 24 subjects, and equal numbers of women and men, is 12. So a difference in average

rank of 3 is one quarter of the way between a completely gender neutral outcome and a complete sorting with men in all the highest ranks.

	PG	*PG w/rank
Women	11.04	10.04
Men	11.96	12.96
std-err, women	1.00	0.94
Std-err, men	1.01	1.03
p-value (2-tailed)	0.52	0.04

Table 3: Rank, scored by number of subjects below in the hierarchy

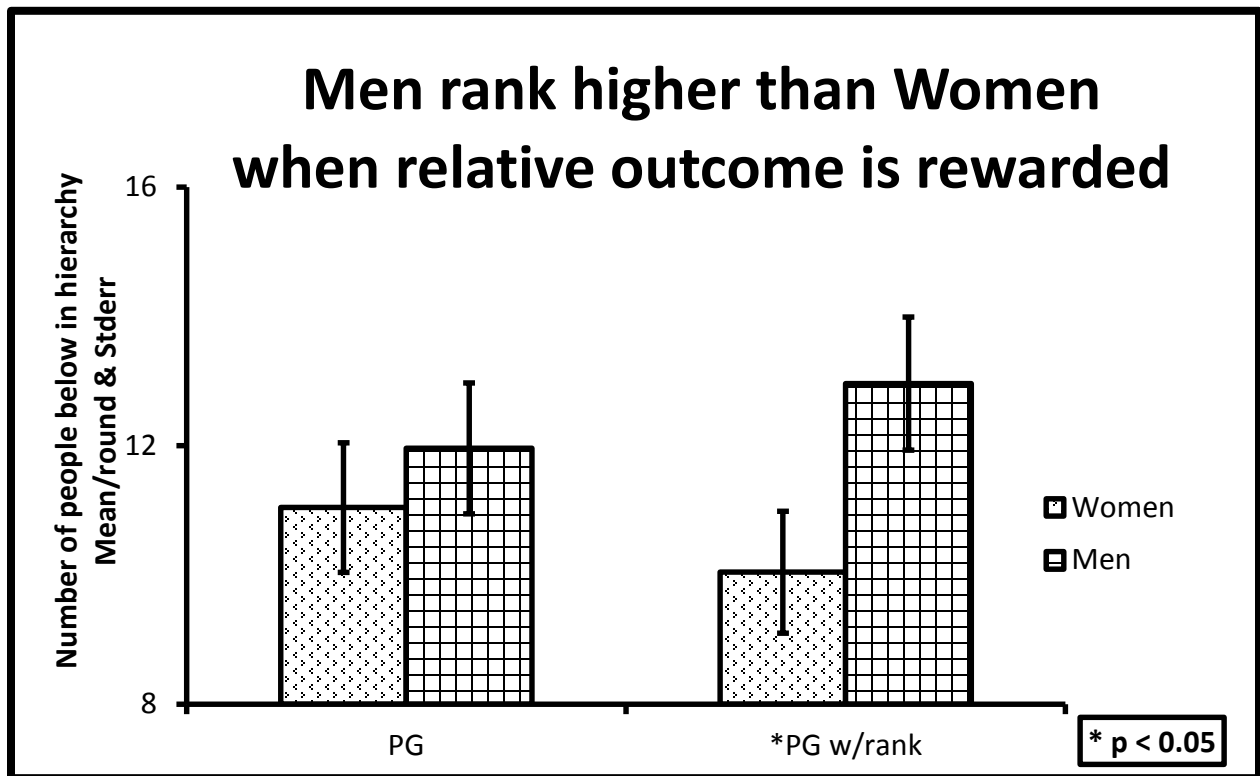


Figure 3: Rank, scored by number of subjects below in the hierarchy

3.4 Costs imposed by male punishment

Men punish more than women in the treatment with rank-based payoffs. Punishment imposes costs on both the punished and the punisher. Each unit of punishment decreases the group payoff by 51 ECU. With rank-based payoffs, men administer 4.4 more units of punishment per round than women. The cost of this extra punishment is 224.4 ECU, which is \$2.24/round per male subject. Moreover, it is possible that male punishment increases female punishment, thus the cost imposed by men on the group may be larger than \$2.24/round per male subject.

4 Discussion

This study reports that men punish more than women in a public goods game with explicit rank payoffs. Men obtain higher rank in this setting and in doing so impose large costs on the others.

This result raises three questions. 1. Is there a robust gender difference in cooperative behavior? 2. What would gender differences in cooperation mean? 3. What other studies are suggested?

Question #1: Is there a robust gender difference in cooperative behavior?

This study reports that men punish at roughly twice the level of women in a situation that includes explicit rank-based payoffs. The result of higher punishment by men is that men obtain higher rank at the cost of women's rank, and the overall group's payoff.

How robust is this conclusion? The treatment where men punish more than women includes two aspects that differ from most prior laboratory experiments. First, the punishment technology uses high-power punishment. Specifically, this study uses 50:1 technology where a punisher suffers one unit of cost for every fifty units imposed on the punished individual. Second, the gender difference in punishment occurs in the presence of explicit rank-based incentives.

Are high-powered punishment and rank-based incentives common? Outside the laboratory, high-powered punishment and rank-based reward may be the norm (Burnham 2014, Burnham 2015). However, in laboratory experiments the most common environment uses lower-powered punishment (e.g. 3:1), and no explicit rank-based payoffs.

Thus, the first set of follow up experiments is to look for gender differences in public goods games with punishment that are more similar to the games commonly played in laboratories. Even if 3:1 punishment without rank-based payoffs is rare outside the laboratory, it is important to connect to the prior literature by running experiments with treatments more similar to the bulk of the prior studies.

This study uses a traditional US college subject pool. Because this set of subjects may be unusual, it would be informative to run similar experiments with subjects drawn from other sorts of subject pools. In particular, it would be interesting to examine gender differences with subjects drawn from populations that exhibit high levels of anti-social punishment. Would the people who administer anti-social punishment being even more extreme with high-powered punishment and with incentives to punish?

It would be interesting to look at situations that allow for counter-punishment. In the current study, individuals can move up in rank by punishing others. With counter-punishment, there would be a way to discipline this rank-seeking punishment. Furthermore, one could examine counter-punish specifically in societies that exhibit anti-social punishment.

Another set of studies would alter the gender ratio of the groups. In this study, every group of 24 subjects was 50% women and 50% men. The identity and gender of the counterparts was not

mentioned to the subjects, but the subjects could see each other, at least in the waiting area before the experiment began. Thus, it is reasonable to expect that the subjects knew the overall group had many men and many women. It would be interesting to look at the relationship between gender composition and behavior.

If people respond to being punished by punishing more in subsequent rounds, and men punish more than women in some settings, perhaps women would punish even less in an all-women setting, and perhaps men would punish even more in an all-men group. If some gender composition effects exist, is there some threshold effect where a few bad apples (e.g., men) spoil the group?

Question 2: What would gender differences in cooperation mean?

There were two motivations for studying gender. The practical motivation is to enhance design of institutions and incentives to promote cooperation. If women and men cooperate differently, then we may need to tailor our cooperative solutions based on gender identity. The idea of gender differences in cooperative system design is analogous to the effort in medicine to tailor therapy based on gender (Geer and Shen 2009, Oertelt-Prigione and Regitz-Zagrosek 2011, Maric-Bilkan and Manigrasso 2012).

The theoretical motivation for this study is to evaluate the claim that strong reciprocity exists in humans. The results of this study may favor the anti-strong reciprocity view, but this conclusion is far from definitive. As far as can be discerned by reading the literature, the proponents of strong reciprocity do make any prediction regarding who will administer the altruistic punishment.

Question 3: What other types of studies are suggested?

Our understanding of cooperation could be improved by examining the relationship between cooperative behavior and any well-reasoned aspect of genotype and phenotype. Some obvious candidates include the 'central six' of cognitive ability, openness, conscientiousness, extraversion, agreeableness, and neuroticism. Any physical characteristic can be studied including height, weight, body-mass, fluctuating asymmetry, beauty, 2D-4D ratio, and hormone levels. Within genotype, a wide range of markers is available for study.

To date, little is known about the sources of individual variation in cooperative behaviors (Camerer 2003). A study of ultimatum game rejections, using monozygotic and polyzygotic twins as experimental subjects, reports over 40% of behavior is explained by genetic effects (Wallace, Cesarini et al. 2007). Ultimatum game rejections are positively correlated with baseline activation in the right prefrontal cortex (Knoch, Gianotti et al. 2010).

One class of study is to correlate cooperative behavior to aspect of genotype or phenotype as has been done for testosterone and the ultimatum game (Burnham 2007). An alternative approach is to examine causation by creating control and test groups that differ in some exogenous factor. One could, for example, administer oxytocin in a controlled experiment as has been done for the trust game and the ultimatum game (Zak, Kurzban et al. 2005, Zak, Kurzban et al. 2009).

Another approach would be to manipulate the information that the participants receive. To elaborate one of myriad possibilities, participants could be shown images of counterparts where the images have been altered, for example, along a dimension of beauty, size, age, or femininity. Such designs reveal information about the functioning of cooperative mechanisms. Performing such experiments requires expertise in both cooperation and the relevant field (e.g., age effects).

In summary, the modern approach to altruism and cooperation dates back at least to W.D. Hamilton's seminal 1964 papers. This paper seeks to improve our understanding of cooperation by examining gender differences. Further studies on gender and other sources of individual variation can help resolve two of the most important issues in cooperation.

First, one size may not fit all; thus, we need to know more about who cooperates, and in what way, in order to produce desired outcomes. Second, did the mechanisms that modulate cooperation arise by selfless group selection as argued by the proponents of strong reciprocity? Answering this question has obvious theoretical ramifications, and is also practical as understanding selective pressures allows for better hypotheses.

Acknowledgments

Helpful comments on experimental design were provided by Dominic Johnson, Robert Kurzban, Nikos Nikiforakis, and Toshio Yamagishi. The z-tree code and the experimental instructions for this experiment were modified versions of those received from Nikos Nikiforakis and Hans-Theo Normann from their 2008 paper 'A Comparative Statics Analysis of Punishment in Public-Good Experiments'. Dominic Donato wrote the new z-tree code, and ran the software during the experiments. The experiments were run in Chapman University's ESI laboratory, under the guidance of Jennifer Cunningham and Jennifer Brady. The final experimental design and manuscript text remains the responsibility of the author.

5 Literature Cited

- Asiedu, E. and M. Ibanez (2014). The weaker sex? Gender differences in punishment across matrilineal and patriarchal societies, GlobalFood Discussion Papers.
- Barankay, I. (2011). "Gender differences in productivity responses to performance rankings: Evidence from a randomized workplace experiment." Unpublished working paper, The Wharton School.
- Binmore, K. (2006). "Why do People Cooperate." Politics, Philosophy and Economics **5**(1): 81-96.
- Brosnan, S. F. and F. B. M. de Waal (2003). "Monkeys reject unequal pay." Nature **425**(18 September): 297-299.
- Brosnan, S. F., H. C. Schiff and F. B. De Waal (2005). "Tolerance for inequity may increase with social closeness in chimpanzees." Proceedings of the Royal Society B: Biological Sciences **272**(1560): 253-258.
- Brosnan, S. F., J. B. Silk, J. Henrich, M. C. Marenco, S. P. Lambeth and S. J. Schapiro (2009). "Chimpanzees (Pan troglodytes) do not develop contingent reciprocity in an experimental task." Animal cognition **12**(4): 587-597.
- Burnham, T. C. (2007). "High-testosterone men reject low ultimatum game offers." Proceedings of the Royal Society (B) **274**: 2327-2330.
- Burnham, T. C. (2014). "Public goods with high-powered punishment: high cooperation and low efficiency." Journal of Bioeconomics **17**(2): 173-187.
- Burnham, T. C. (2015). "Public goods with punishment & payment for relative rank." SSRN working paper # 2623763: 1-17.
- Burnham, T. C. and D. Johnson (2005). "The Biological and Evolutionary Logic of Human Cooperation." Analyse & Kritik **27**: 113-135.
- Burton-Chellew, M. N. and S. A. West (2013). "Prosocial preferences do not explain human cooperation in public-goods games." Proceedings of the National Academy of Sciences **110**(1): 216-221.
- Camerer, C. (2003). Behavioral Game Theory: Experiments in Strategic Interaction. Princeton, Princeton University Press.
- Clutton-Brock, T. H. and G. A. Parker (1995). "Punishment in Animal Societies." Nature **373**: 209-216.
- Cronin, K. A., K. K. Schroeder and C. T. Snowdon (2010). "Prosocial behaviour emerges independent of reciprocity in cottontop tamarins." Proceedings of the Royal Society B: Biological Sciences **277**(1701): 3845-3851.
- Croson, R. and U. Gneezy (2009). "Gender differences in preferences." Journal of economic literature: 448-474.
- Dreber, A., D. Rand, D. Fudenberg and M. Nowak (2008). "Winners don't punish." Nature **452**(20 March 2008): 348-351.

Eckel, C. C. and P. J. Grossman (2001). "Chivalry and solidarity in ultimatum games." Economic Inquiry **39**(2): 171-188.

Eckel, C. C. and P. J. Grossman (2008). "Differences in the economic decisions of men and women: Experimental evidence." Handbook of experimental economics results **1**: 509-519.

Fehr, E., U. Fischbacher and S. Gächter (2002). "Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms." Human Nature **13**: 1-25.

Fehr, E. and S. Gächter (2000). "Cooperation and Punishment in Public Goods Experiments." American Economic Review **90**: 980-994.

Fehr, E. and S. Gächter (2002). "Altruistic Punishment in Humans." Nature **415**: 137-140.

Fowler, J. H. (2005). "Altruistic punishment and the origin of cooperation." PNAS **102**(19): 7047-7049.

Geer, E. B. and W. Shen (2009). "Gender differences in insulin resistance, body composition, and energy balance." Gender medicine **6**: 60-75.

Gintis, H. (2000). "Strong Reciprocity and Human Sociality." Journal of Theoretical Biology **206**: 169-179.

Gneezy, U., K. L. Leonard and J. A. List (2008). Gender differences in competition: Evidence from a matrilineal and a patriarchal society, National Bureau of Economic Research.

Hamilton, W. D. (1964). "The Genetical Evolution of Social Behavior I and II." Journal of Theoretical Biology **7**(1): 1-16, 17-52.

Henrich, J., S. J. Heine and A. Norenzayan (2010). "The Weirdest People in the World?" Behavioral and Brain Sciences **33**: 61-135.

Herrmann, B., C. Thöni and S. Gächter (2008). "Antisocial Punishment Across Societies." Science **319**(5868): 1362-1367.

Jensen, K., J. Call and M. Tomasello (2007). "Chimpanzees Are Rational in an Ultimatum Game." Science **318**(5 October): 107-109.

Jensen, K., J. Call and M. Tomasello (2007). "Chimpanzees are vengeful but not spiteful." Proceedings of the National Academy of Sciences **104**(32): 13046-13050.

Jensen, K., B. Hare, J. Call and M. Tomasello (2006). "What's in it for me? Self-regard precludes altruism and spite in chimpanzees." Proceedings of the Royal Society B: Biological Sciences **273**(1589): 1013-1021.

Johnson, D., P. Stopka and S. Knights (2003). "The Puzzle of Human Cooperation." Nature **421**(27 February): 911-912.

Knoch, D., L. R. R. Gianotti, T. Baumgartner and E. Fehr (2010). "A Neural Marker of Costly Punishment Behavior." Psychological science.

Krasnow, M. M., A. W. Delton, L. Cosmides and J. Tooby (2015). "Group Cooperation without Group Selection: Modest Punishment Can Recruit Much Cooperation."

Kurzban, R. and P. DeScioli (2013). "Adaptationist punishment in humans." Journal of Bioeconomics: 1-11.

Maric-Bilkan, C. and M. B. Manigrasso (2012). "Sex differences in hypertension: contribution of the renin-angiotensin system." Gender medicine **9**(4): 287-291.

Mazur, A. (2015). A Biosocial Model of Status in Face-To-Face Groups. Evolutionary Perspectives on Social Psychology, Springer: 303-315.

Mazur, A. and A. Booth (1998). "Testosterone and Dominance in Men." Behavioral and Brain Sciences **21**: 353-363.

Molina, J. A., J. I. Giménez-Nadal, J. A. Cuesta, C. Gracia-Lazaro, Y. Moreno and A. Sanchez (2013). "Gender differences in cooperation: experimental evidence on high school students."

Niederle, M. and L. Vesterlund (2007). "Do women shy away from competition? Do men compete too much? ." Quarterly Journal of Economics **122**(3).

Nowak, M. and K. Sigmund (1998). "Evolution of Indirect reciprocity by image scoring." Nature **393**: 573-577.

Oertelt-Prigione, S. and V. Regitz-Zagrosek (2011). Sex and gender aspects in clinical medicine, Springer Science & Business Media.

Price, M., L. Cosmides and J. Tooby (2002). "Punitive Sentiment as an anti-free rider psychological device." Evolution and Human Behavior **23**: 203-231.

Proctor, D., R. A. Williamson, F. B. M. de Waal and S. F. Brosnan (2013). "Chimpanzees play the ultimatum game." Proceedings of the National Academy of Sciences **110**(6): 2070-2075.

Raihani, N. J., A. Thornton and R. Bshary (2012). "Punishment and cooperation in nature." Trends in Ecology & Evolution **27**(5): 288-295.

Rand, D. G. and M. A. Nowak (2011). "The evolution of antisocial punishment in optional public goods games." Nature Communications **2**: 434.

Riedl, K., K. Jensen, J. Call and M. Tomasello (2012). "No third-party punishment in chimpanzees." Proceedings of the National Academy of Sciences **109**(37): 14824-14829.

Rutte, C. and M. Taborsky (2007). "Generalized reciprocity in rats." PLoS biology **5**(7): e196.

Sapolsky, R. M. (2005). "The influence of social hierarchy on primate health." Science **308**(5722): 648-652.

Selten, R. and A. Ockenfels (1998). "An experimental solidarity game." Journal of Economic Behavior & Organization **34**(4): 517-539.

Shinada, M., T. Yamagishi and Y. Ohmura (2004). "False friends are worse than bitter enemies: "altruistic" punishment of in-group members." Evolution and Human Behavior **25**: 379-393.

Sigmund, K. (2007). "Punish or perish? Retaliation and collaboration among humans." Trends in Ecology & Evolution **22**(11): 593-600.

Silk, J. B., S. F. Brosnan, J. Vonk, J. Henrich, D. J. Povinelli, A. F. Richardson, S. P. Lambeth, J. Mascaró and S. J. Schapiro (2005). "Chimpanzees are indifferent to the welfare of other group members." Nature **435**: 1357-1359.

Singer, T., B. Seymour, J. P. O'Doherty, K. E. Stephan, R. J. Dolan and C. D. Frith (2006). "Empathic neural responses are modulated by the perceived fairness of others." Nature **439**(7075): 466-469.

Solnick, S. J. (2001). "Gender differences in the ultimatum game." Economic Inquiry **39**(2): 189-200.

Talbot, C. F., H. D. Freeman, L. E. Williams and S. F. Brosnan (2011). "Squirrel monkeys' response to inequitable outcomes indicates a behavioural convergence within the primates." Biology letters **7**(5): 680-682.

Tran, A. and R. Zeckhauser (2012). "Rank as an inherent incentive: Evidence from a field experiment." Journal of Public Economics **96**(9): 645-650.

Trivers, R. L. (1971). "The Evolution of Reciprocal Altruism." Quarterly Review of Biology **46**(1): 35-57.

Trivers, R. L. (2004). "Mutual Benefit at All Levels of Life." Science **304**(14 May 2004): 964.

Van Leeuwen, E. J., E. Zimmermann and M. D. Ross (2011). "Responding to inequities: Gorillas try to maintain their competitive advantage during play fights." Biology letters **7**(1): 39-42.

Vonk, J., S. F. Brosnan, J. B. Silk, J. Henrich, A. S. Richardson, S. P. Lambeth, S. J. Schapiro and D. J. Povinelli (2008). "Chimpanzees do not take advantage of very low cost opportunities to deliver food to unrelated group members." Animal Behaviour **75**(5): 1757-1770.

Wallace, B., D. Cesarini, P. Lichtenstein and M. Johannesson (2007). "Heritability of ultimatum game responder behavior." PNAS **104**(40): 15631-15634.

West, S. A., C. E. Mouden and A. Gardner (2011). "Sixteen common misconceptions about the evolution of cooperation in humans." Evolution and Human Behavior **32**: 231-262.

Wiessner, P. (2005). "Norm enforcement among the Ju/'hoansi Bushmen." Human Nature **16**(2): 115-145.

Wilson, D. S. (1975). "A theory of group selection." Proceedings of the National Academy of Sciences **72**(1): 143-146.

Yamagishi, T. (1986). "The Provision of a Sanctioning System as a Public Good." Journal of Personality and Social Psychology **51**(1): 110-116.

Yee, J. R., S. A. Cavigelli, B. Delgado and M. K. McClintock (2008). "Reciprocal affiliation among adolescent rats during a mild group stressor predicts mammary tumors and lifespan." Psychosomatic medicine **70**(9): 1050-1059.

Zahavi, A. (1975). "Mate Selection--A selection for a Handicap." Journal of Theoretical Biology **53**: 205-214.

Zak, P. J., R. Kurzban, S. Ahmadi, R. S. Swerdloff, J. Park, L. Efremidze, K. Redwine, K. Morgan and W. Matzner (2009). "Testosterone Administration Decreases Generosity in the Ultimatum Game." PLoS ONE **4**(12): e8330.

Zak, P. J., R. Kurzban and W. T. Matzner (2005). "Oxytocin is associated with human trustworthiness." Hormones and Behavior **48**: 522-527.