

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

Terence C. Burnham<sup>1</sup>

## Abstract

A laboratory experiment reports on gender, cooperation, and punishment in two public goods games using high-powered punishment. In a public goods games with punishment, no statistically significant differences between men and women are reported. In a modified game that includes 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 women. These results contribute to the debate about the origins and maintenance of cooperation.

## Keywords

cooperation, reciprocity, punishment, public goods, altruism, gender

## Introduction

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

This paper seeks to make two contributions to the study of cooperation. First, it reports gender differences in punishment behavior. Second, these gender differences may inform the ongoing and important debate regarding the existence of “strong reciprocity.”

The first motivation for this work is to study gender differences in cooperation. The world is filled with organizations that include both men and women. Indeed, single-sex institutions are the exception. Furthermore, many such situations include the ability to punish others. These punishments range from loss of reputation to direct financial impacts, including loss of employment.

Although mixed gender situations with the ability to punish are the norm outside the laboratory, gender studies of cooperation and punishment are relatively rare. Furthermore, existing studies reveal no clear relationship between gender and certain cooperative behaviors, reporting rather that men are more cooperative than women in some settings (Balliet et al. 2011). However, 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’ dilemmas and social dilemma games (Croson and Gneezy 2009).

While existing studies are ambiguous, 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 et al. 2006), and punishment is reported to vary between patriarchal and matriarchal societies (Asiedu and Ibanez 2014). More broadly, men are reported to be less cooperative (Molina et al. 2013) and less generous (Selten and Ockenfels 1998).

Gender effects in rejection behavior in the ultimatum game are reported to vary based on the setting 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

<sup>1</sup>Chapman University, Orange, CA, USA

## Corresponding Author:

Terence C. Burnham, Associate Professor, Chapman University, One University Avenue, Orange, CA 92866-1005, USA.  
 Email: [terry@post.harvard.edu](mailto:terry@post.harvard.edu)



reported to reject more than men (Solnick 2001). Presumably, there was little or no possibility of physical violence within either of these experiments, but the perception of such risk may be higher in face-to-face interactions.

Second, status impacts cooperative behavior (Eckel, Fatas, and Wilson 2010), and women may have different feelings regarding status and rank than men. If so, punishment may be a tool used by some individuals to advance in relative rank (Burnham 2015). When given a choice, women chose a competitive tournament setting less than men (Niederle and Vesterlund 2007), although in some cultures, women are reported to behave more competitively than men (Gneezy, Leonard, and List 2009). Both men and women reveal a preference for rank (Tran and Zeckhauser 2012), while in a field experiment, women are less “heedful” of rank than men (Barankay 2011).

The first goal of this current study is to contribute to the understudied area of gender and costly cooperation. This is an important area, and the previous research is incomplete and inconclusive.

The second motivation is that gender differences in cooperation may inform the debate regarding strong reciprocity. Strong reciprocity is an idea where individuals sacrifice to help other people (Fehr, Fischbacher, and Gächter 2002; Gintis 2000). The term *strong reciprocity* was created to explain certain costly behaviors that help the group at the expense of the actor who chooses the behavior.

One strand of the strong reciprocity literature examines “altruistic punishment.” In some settings, punishment can increase cooperation (Fehr and Gächter 2002; Yamagishi 1986). Costly punishment is puzzling in situations where the punisher cannot be repaid. The proponents of strong reciprocity view punishment in settings where repayment is impossible as the altruistic manifestation of an altruistic, group-selected human nature (Fehr and Gächter 2002).

While agreeing on the experimental results, critics of strong reciprocity question the group selection foundation for these behaviors (Binmore 2006; Burnham and Johnson 2005; Burton-Chellew and West 2013; Dreber et al. 2008; Johnson, Stopka, and Knights 2003; Krasnow et al. 2015; Kurzban and DeScioli 2013; Price, Cosmides, and Tooby 2002; Trivers 2004; West, Mouden, and Gardner 2011). One idea that unifies some of the criticism of strong reciprocity is that mechanisms that were forged by individual selection can benefit the group in particular settings (Burnham and Johnson 2005; Johnson et al. 2003; Trivers 2004). Therefore, it is not possible to determine the ultimate cause of a behavior by observing its impact in one setting, particularly an evolutionarily novel setting such as an experimental laboratory.

The debate regarding strong reciprocity remains important and unresolved. One method to inform the debate is to examine the characteristics of punishers and evaluate whether punishers share characteristics that are associated with altruism. In this study, for example, we examine gender and punishment. If punishment is altruistic and women were

more generous than men, we might expect women to punish more. Conversely, if we were to find that women punish less than men, we might question whether punishment is altruistic.

This study includes a treatment with monetary payment based in part on relative performance. This treatment is included because almost every area of human endeavor is based at least in part on relative performance (Burnham 2017). For example, grades are usually based on relative performance, and grades play a central role in college admissions and the labor market generally. Once in the labor market, up to 20 percent of U.S. businesses use some version of a forced evaluation system based on relative performance (Bates 2003; Gary 2001). Sports are almost entirely based on relative performance; this year’s champion is better than the competitors. Similarly, elections choose the candidate with the most votes.

This experiment examines gender differences in punishment in public goods games—one with explicit monetary incentives based on relative performance and a second without such monetary incentives.

## 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.

There is a baseline treatment and a rank-based treatment. The baseline game is a public goods game based on Fehr and Gächter’s (2002) “perfect stranger” design. One important modification is that while Fehr and Gächter inflict three units of punishment for each unit paid to punish (3:1 hereafter), the baseline in this experiment uses “high-powered punishment” that inflicts 50 units of monetary punishment for each 1 unit of cost to the punishment (50:1 hereafter).

Subjects in each session played public goods games with and without rank-based payoffs (Burnham 2015). The order of the treatments was counterbalanced between sessions so that half the subjects played the public goods game with rank-based payoffs first.

In each public goods game, the 24 subjects were allocated to six groups of four subjects. Groups were rematched 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 the current period. As noted, each session had equal numbers of women and men.

Subjects were given written instructions that explained the structure of the game, the composition of groups in each period, and 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. To participate, each subject was required to get all the payoff examples correct.

The subjects sat in four rows with six individuals per row. All decisions were made via computers, and each subject had his or her own computer. Three-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 of the public goods games, 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 four 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 their contribution to the public good. Each player could allocate up to 10 units of punishment to each of the three other group members. A unit of punishment cost the punisher 1 ECU and reduced the punished player's payoff by 50 ECU.

The baseline condition as described previously is one of the two treatments. It is Fehr and Gächter's (2002) perfect stranger treatment using a 50:1 punishment technology. The second treatment uses the baseline structure with a single modification of adding a set of rank-based rewards. The rank payoffs were designed to 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.

The specific rank payoffs were determined as follows (and shown in Table 1). At the end of the public goods game with punishment and rank-based payoffs, the 24 subjects in the session were ranked from 1st to 24th. First place was given a bonus of 550 ECUs. Second place earned 500 ECUs. Payoffs decreased by 50 ECUs per place except that subjects in places 12 and 13 each earned 0 ECUs. The subjects were provided the following wording in addition to a table with all 24 rows for rank and payoff: "After this experiment is over, you will be ranked from 1st to 24th place. There are financial rewards and penalties based on where you rank. Specifically, the table below explains how your earnings from the experiment will be adjusted based on your relative rank."

The 50:1 punishment technology allows for significant negative payouts so that the total amount of punishment can exceed the endowment plus accrued earnings. After each round, the total balance for each subject is calculated as endowment plus earnings (minus losses) from prior rounds, plus earnings (minus losses) from this round. If this total became negative for any player, then all the play would be

**Table 1.** Explicit Payoffs to Relative Rank.

Place	Adjustment in Experimental Currency Units
1st place	+550
2nd place	+500
...	...
11th place	+50
12th place	0
13th place	0
14th place	-50
...	...
24th place	-550

**Table 2.** Contributions to the Public Good (PG) by Treatment and Gender.

	PG	PG with Rank
Women	17.03	18.00
Men	17.98	18.64
SE, women	.42	.42
SE, men	.39	.47
<i>p</i> value (two-tailed)	.10	.32

stopped. The endowment was large (1,500 or 2,000 ECUs) relative to the public good payoffs to provide a buffer against punishment costs. The stopping rule was explained to the subjects in advance.

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 (4 pm) on a Tuesday, Wednesday, or Thursday.

No deception was used in the experiment.

## Results

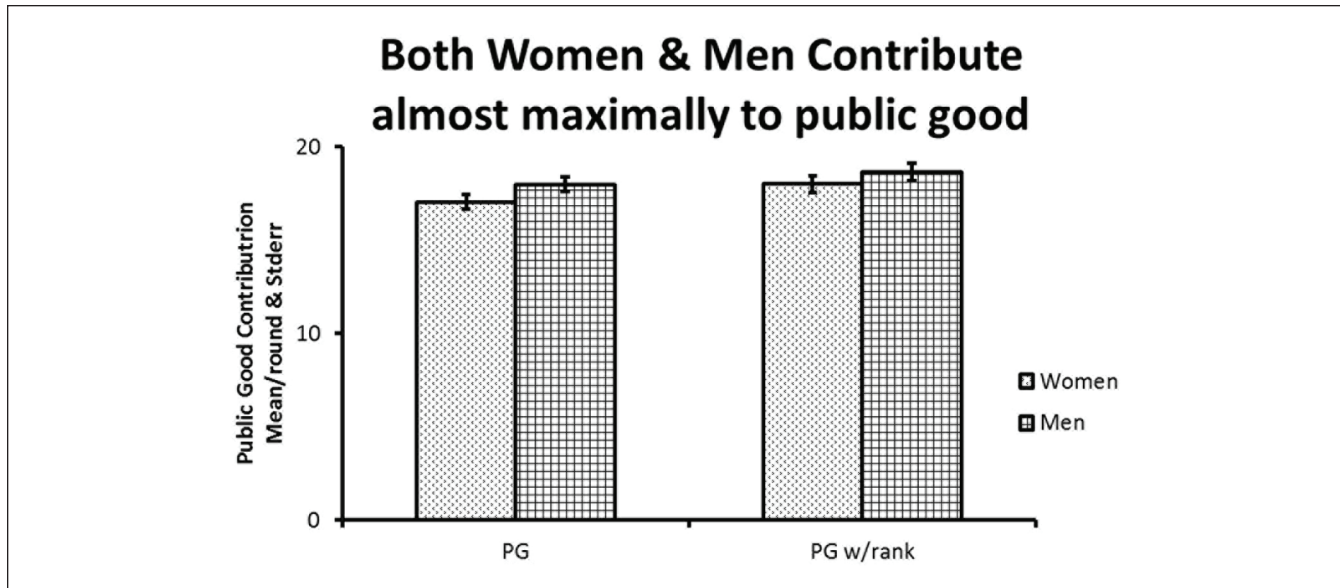
### Contributions

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 2, 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 (SPSS, independent samples *t* test).

### Punishment

There is no statistically significant difference in punishment by men and women in the standard public goods game (Table 3, Figure 2). Women administer 1.76 units punishment per round



**Figure 1.** Contributions to the public good by treatment and gender.

**Table 3.** Punishment Administered by Treatment and Gender.

	Public Good	Public Good with Rank
Women	1.76	4.34
Men	2.46	8.74
SE, women	.42	.68
SE, men	.50	1.12
<i>p</i> value (two-tailed)	.29	.00

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 of 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  $p < .01$  level (SPSS, independent samples *t* test).

In summary, rank-based incentives caused subjects to punish more. Male subjects were particularly apt to punish more in the rank-based condition.

## Rank

In the treatment with rank-based payoffs, after the end of the public goods game, subjects are ranked from 1st to 24th based on earning in that public goods game. First place is the best, and 24th place is the worst.

For the numerical analysis that follows, we felt it would be more intuitive to use a ranking system where a higher

number means a higher rank. Thus, we convert the subject rankings from the aforementioned in the following analysis so that a ranking of 23 is the highest possible and means that the 23 other subjects ranked below this individual. This ranking system has the desired feature that a larger number equates to a 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 (Table 4, Figure 3). Women obtain an average rank of 10.04, almost three full places below the 12.96 obtained by men. This gender difference is significant at the  $p < .05$  level (SPSS, independent samples *t* test).

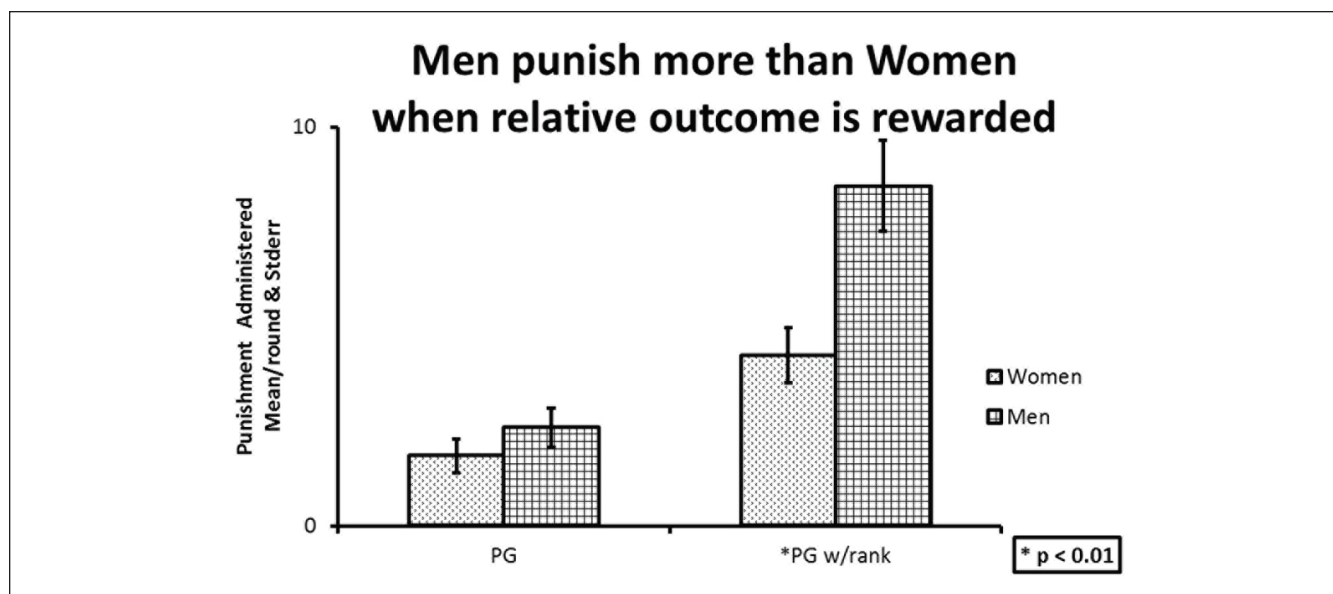
This difference in rank of almost 3 slots between men and women is large. The maximum possible difference with 24 subjects and equal numbers of women and men is 12. So a difference in average rank of 3 is one-quarter of the way between a completely gender-neutral outcome and a complete sorting with men in all the highest ranks.

## Punishment of Full Cooperators

How do men achieve their higher rank in the public goods game with pay for relative performance? Part of the explanation is that men punish full cooperators more than women.

A full cooperator can be defined as someone who contributes 100% to the public good. Punishment of full cooperators cannot be designed to induce higher cooperation because more cooperation is not possible. However, in the context of relative performance, punishing full cooperators will move the punisher up in rank. This is because the cost of punishment in





**Figure 2.** Punishment administered by treatment and gender.

**Table 4.** Rank, Scored by Number of Subjects below in the Hierarchy.

	Public Good	Public Good with Rank
Women	11.04	10.04
Men	11.96	12.96
SE, women	1.00	.94
SE, men	1.01	1.03
p value (two-tailed)	.52	.04

this (and many) public goods games is greater for the punished than the punisher.

Table 5 and Figure 4 summarize the punishment of full cooperators by gender and treatment. There is no statistically significant difference in punishment of full cooperators by men and women in the standard public goods game. Women administer .77 units of punishment per round to full cooperators while men administer an average of .91 punishment units to full cooperators.

In contrast, in the presence of rank-based payoffs, men administer almost exactly twice as much punishment to full cooperators. Women administer an average of 3.27 units of punishment per round to full cooperators while men administer an average of 6.54 punishment units per round to full cooperators. This difference is significant at the  $p < .01$  level (SPSS, independent samples  $t$  test).

### 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 in subsequent games, thus the total cost imposed by men on the group may be larger than \$2.24/round per male subject.

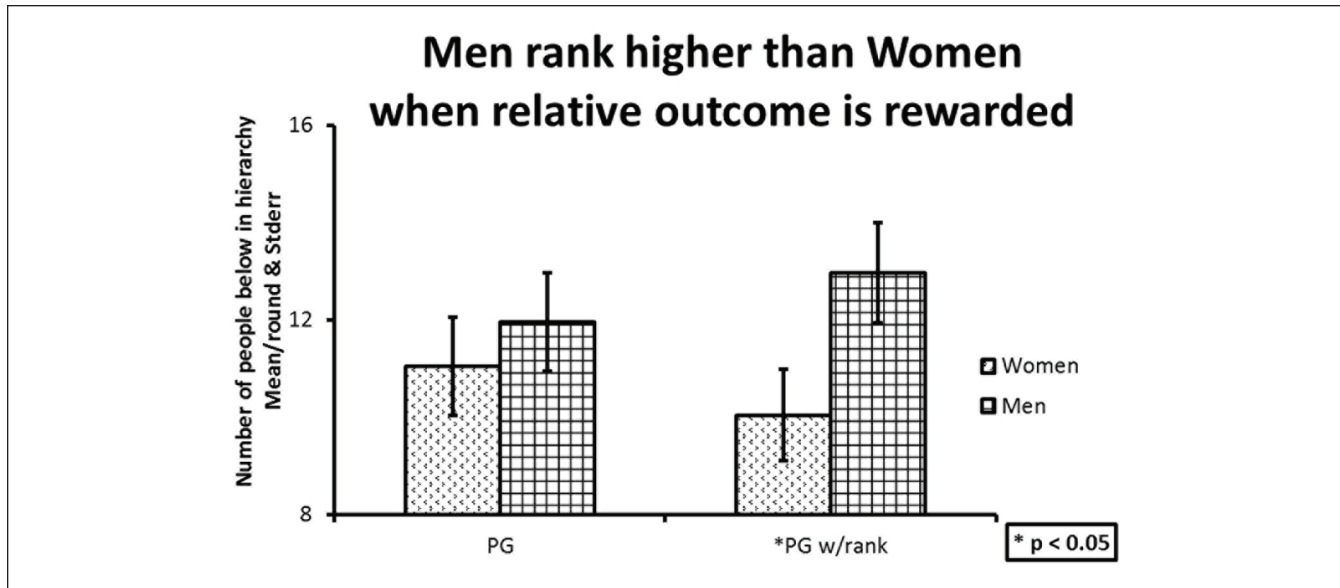
### Discussion

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 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 outcomes, and punishment reduces the group's payoff. Furthermore, men punished full cooperators more than women in the context of rank-based payoffs.

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

Outside the laboratory, high-powered punishment and rank-based reward may be the norm (Burnham 2015, 2017). However, in laboratory experiments, the most common environment uses lower powered punishment (e.g., 3:1) and no explicit rank-based payoffs. Thus, even if 3:1 punishment



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

**Table 5.** Punishment Administered to Full Cooperators by Treatment and Gender.

	Public Good	Public Good with Rank
Women	.77	3.27
Men	.91	6.54
SE, women	.26	.61
SE, men	.25	1.10
p value (two-tailed)	.50	.00

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.

Another set of potentially interesting studies could alter the gender ratio of the groups. In this study, every group of 24 subjects was 50 percent women and 50 percent 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.

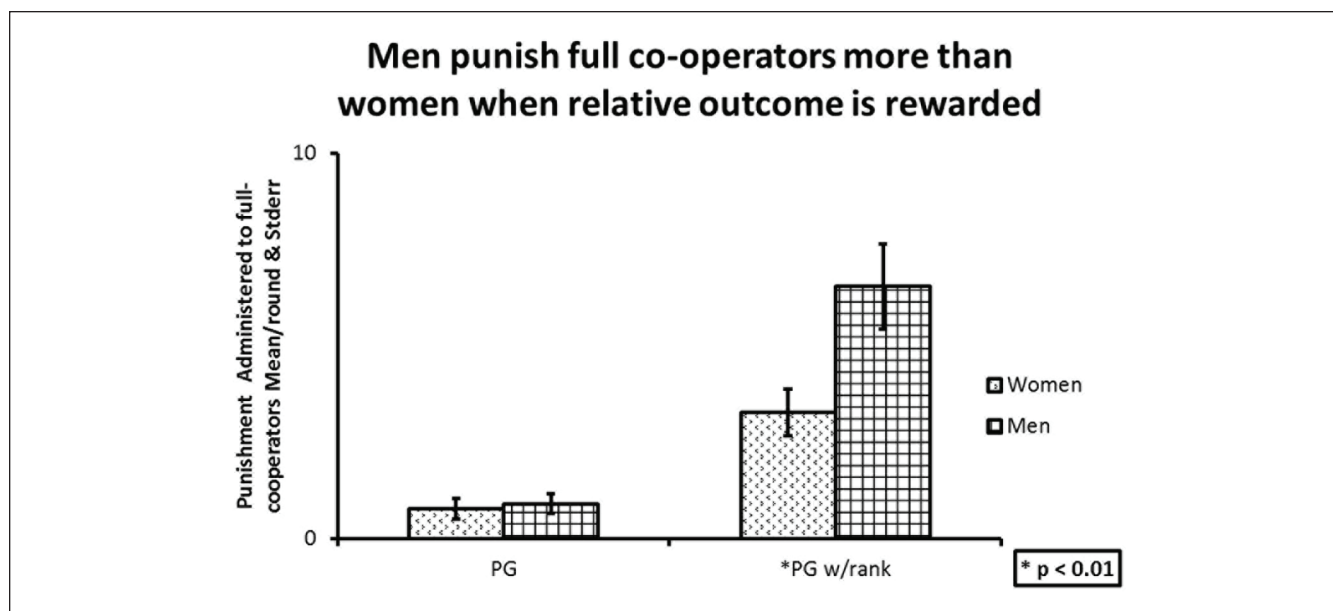
There were two motivations for studying gender. A practical motivation is to enhance design of institutions and incentives to promote cooperation. If women and men modulate

cooperation 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; Maric-Bilkan and Manigrasso 2012; Oertelt-Prigione and Regitz-Zagrosek 2011).

A different motivation for this study is to evaluate the claim that strong reciprocity exists in humans. The proponents of strong reciprocity interpret punishment as a form of altruism. In this study, men use punishment to advance their own interests at the cost of others. This form of punishment is not altruistic.

Fehr and Gächter (2002) argue that although punishment can deter free-riding, punishment is itself subject to a free-rider problem. Thus, when we see punishment in one-shot situations, such punishment should be categorized as altruistic because the punishing individual provides a public good by paying a cost on behalf of the group to deter defection. An alternative to altruism is that participants are attending to their local rank even when rank is not part of an explicit monetary reward system. The results of this study add to the growing literature that questions the alleged altruistic intention of punishment (Burnham and Johnson 2005; West et al. 2011; Yamagishi et al. 2012).

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 to produce desired outcomes. Second, did the mechanisms that modulate cooperation arise by selfless group selection as argued by the proponents of strong reciprocity?



**Figure 4.** Punishment administered to full cooperators by treatment and gender.

### 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 Economics Science Institute laboratory under the guidance of Jennifer Cunningham and Jennifer Brady. The paper was improved by the comments of the editor and anonymous referees. The final experimental design and manuscript text remain the responsibility of the author.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The experiments were funded by Chapman University.

### References

- Asiedu, Edward, and Marcel Ibanez. 2014. “The Weaker Sex? Gender Differences in Punishment across Matrilineal and Patriarchal Societies.” Retrieved November 1, 2017 (<http://purl.umn.edu/165743>).
- Balliet, Daniel, Norman P. Li, Shane J. Macfarlan, and Mark Van Vugt. 2011. “Sex Differences in Cooperation: A Meta-analytic Review of Social Dilemmas.” *Psychological Bulletin* 137(6):881–909.
- Barankay, Iwan. 2011. “Gender Differences in Productivity Responses to Performance Rankings: Evidence from a Randomized Workplace Experiment.” Unpublished working paper, The Wharton School.
- Bates, Steve. 2003. “Forced Ranking.” *HR Magazine*. Retrieved November 1, 2017 (<https://www.shrm.org/hr-today/news/hr-magazine/pages/0603bates.aspx>).
- Binmore, Ken. 2006. “Why Do People Cooperate.” *Politics, Philosophy and Economics* 5(1):81–96.
- Burnham, Terence C. 2015. “Public Goods with High-powered Punishment: High Cooperation and Low Efficiency.” *Journal of Bioeconomics* 17(2):173–87.
- Burnham, Terence C. 2017. “Public Goods with Punishment & Payment for Relative Rank.” *Psychology and Behavioral Science International Journal* 3(1):1–7.
- Burnham, Terence C., and Dominic Johnson. 2005. “The Biological and Evolutionary Logic of Human Cooperation.” *Analyse & Kritik* 27:113–35.
- Burton-Chellew, Maxwell N., and Stuart A. West. 2013. “Prosocial Preferences Do Not Explain Human Cooperation in Public-goods Games.” *Proceedings of the National Academy of Sciences* 110(1):216–21.
- Clutton-Brock, T. H., and G. A. Parker. 1995. “Punishment in Animal Societies.” *Nature* 373:209–16.
- Croson, Rachel, and Uri Gneezy. 2009. “Gender Differences in Preferences.” *Journal of Economic Literature* 47:448–74.
- Dreber, Anna, David Rand, Drew Fudenberg, and Martin Nowak. 2008. “Winners Don’t Punish.” *Nature* 452:348–51.
- Eckel, Catherine C., Enrique Fatas, and Rick Wilson. 2010. “Cooperation and Status in Organizations.” *Journal of Public Economic Theory* 12(4):737–62.
- Eckel, Catherine C., and Philip J. Grossman. 2001. “Chivalry and Solidarity in Ultimatum Games.” *Economic Inquiry* 39(2):171–88.
- Eckel, Catherine C., and Philip J. Grossman. 2008. “Differences in the Economic Decisions of Men and Women: Experimental Evidence.” *Handbook of Experimental Economics Results* 1:509–19.

- Fehr, Ernst, Urs Fischbacher, and Simon Gächter. 2002. "Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms." *Human Nature* 13:1–25.
- Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90:980–94.
- Fehr, Ernst, and Simon Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415:137–40.
- Fowler, James H. 2005. "Altruistic Punishment and the Origin of Cooperation." *PNAS* 102(19):7047–49.
- Gary, Loren. 2001. "The Controversial Practice of Forced Ranking." Retrieved November 1, 2017 (<https://hbr.org/product/controversial-practice-of-forced-ranking/U0110A-PDF-ENG>).
- Geer, Eliza B., and Wei Shen. 2009. "Gender Differences in Insulin Resistance, Body Composition, and Energy Balance." *Gender Medicine* 6:60–75.
- Gintis, Herbert. 2000. "Strong Reciprocity and Human Sociality." *Journal of Theoretical Biology* 206:169–79.
- Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society." *Econometrica* 77:1637–64.
- Johnson, Dominic, Pavel Stopka, and Stephen Knights. 2003. "The Puzzle of Human Co-operation." *Nature* 421:911–12.
- Krasnow, Max M., Andrew W. Delton, Leda Cosmides, and John Tooby. 2015. "Group Cooperation without Group Selection: Modest Punishment Can Recruit Much Cooperation." *PLoS ONE* 10(4):e0124561.
- Kurzban, Robert, and Peter DeScioli. 2013. "Adaptationist Punishment in Humans." *Journal of Bioeconomics* 15:269–79.
- Maric-Bilkan, Christine, and Michael B. Manigrasso. 2012. "Sex Differences in Hypertension: Contribution of the Renin-angiotensin System." *Gender Medicine* 9(4):287–91.
- Molina, J. Alberto, J. Ignacio Giménez-Nadal, Jose A. Cuesta, Carlos Gracia-Lazaro, Yamir Moreno, and Angel Sanchez. 2013. "Gender Differences in Cooperation: Experimental Evidence on High School Students." *PLoS ONE* 8(12):e83700.
- Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics* 122(3):1067–101.
- Oertelt-Prigione, Sabine, and Vera Regitz-Zagrosek. 2011. *Sex and Gender Aspects in Clinical Medicine*. New York: Springer Science & Business Media.
- Price, Michael, Leda Cosmides, and John Tooby. 2002. "Punitive Sentiment as an Anti-free rider Psychological Device." *Evolution and Human Behavior* 23:203–31.
- Selten, Reinhold, and Axel Ockenfels. 1998. "An Experimental Solidarity Game." *Journal of Economic Behavior & Organization* 34(4):517–39.
- Sigmund, Karl. 2007. "Punish or Perish? Retaliation and Collaboration among Humans." *Trends in Ecology & Evolution* 22(11):593–600.
- Singer, Tania, Ben Seymour, John P. O'Doherty, Klaas E. Stephan, Raymond J. Dolan, and Chris D. Frith. 2006. "Empathic Neural Responses Are Modulated by the Perceived Fairness of Others." *Nature* 439(7075):466–69.
- Solnick, S. J. 2001. "Gender Differences in the Ultimatum Game." *Economic Inquiry* 39(2):189–200.
- Tran, Anh, and Richard Zeckhauser. 2012. "Rank as an Inherent Incentive: Evidence from a Field Experiment." *Journal of Public Economics* 96(9):645–50.
- Trivers, Robert L. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology* 46(1):35–57.
- Trivers, Robert L. 2004. "Mutual Benefit at All Levels of Life." *Science* 304(14):964.
- West, Stuart A., Claire El Mouden, and Andy Gardner. 2011. "Sixteen Common Misconceptions about the Evolution of Cooperation in Humans." *Evolution and Human Behavior* 32:231–62.
- Yamagishi, Toshio. 1986. "The Provision of a Sanctioning System as a Public Good." *Journal of Personality and Social Psychology* 51(1):110–16.
- Yamagishi, Toshio, Yutaka Horito, Nobuhiro Mifune, Hirofumi Hashimoto, Yang Li, Mizuho Shinada, Arisa Miura, Keigo Inukai, Haruto Takagishi, and Dora Simunovic. 2012. "Rejection of Unfair Offers in the Ultimatum Game Is No Evidence of Strong Reciprocity." *Proceedings of the National Academy of Sciences* 109(50):20364–68.

## Author Biography

**Terence C. Burnham** is an economist. His research focuses on the biological causes of human economic behavior. He is an associate professor at Chapman University in Orange, California.