

SCIENTIFIC REPORTS



OPEN

Sex Affects the Relationship Between Third Party Punishment and Cooperation

Claudia Rodriguez-Ruiz¹, José Antonio Muñoz-Reyes², Marta Iglesias-Julios³, Santiago Sanchez-Pages⁴ & Enrique Turiegano¹

Prosocial third-party punishment (3PP) is a punitive behavior against antisocial individuals, which might explain extended cooperativeness in humans. 3PP shows sexual dimorphism, being more frequent in men than in women. We studied whether sexually dimorphic features related to sexual hormones during development (facial dimorphism and 2D:4D) influence the tendency to engage in 3PP in a sample of 511 women and 328 men. After playing a Prisoner's Dilemma, participants had to decide whether to penalize the defection of a third player who had exploited his/her counterpart's cooperation. In line with previous studies, we observe that men are more prone to engage in 3PP than women. We find that this sex difference is due to cooperative men being more likely to punish than cooperative women. In addition, men with higher facial masculinity are less likely to engage in 3PP, whereas no features influence 3PP in women. We discuss the possibility that sex differences in the motivations and fitness implications underlying 3PP might be driving the observed results.

Third party punishment (3PP) is a behavior costly to the individual exerting it and aimed at penalizing individuals who have not directly interacted with the punisher. It is considered to be altruistic when individuals exert it to penalize others' harmful behavior¹. Altruistic 3PP appears in many human societies^{2–4}. Because of that, it has been proposed as one among the possible explanations for extended human cooperation^{1,5,6}, as individuals who engage in altruistic 3PP pay currently a cost to increase social welfare. Still, for altruistic 3PP to be selected for, it must improve the inclusive fitness of individuals exerting it⁷; either directly (i.e. increasing reproductive success through processes linked to reciprocity or reputation) or indirectly^{8–11}. In fact, the attainment of such fitness benefits must be the ultimate cause of prosocial behavior^{7,12,13}, independently of whether the individual receives in exchange future social advantages or an improvement in resources. From this viewpoint, altruistic 3PP is expected to share proximal and evolutionary motivations and influences with other forms of prosociality. Individuals who engage in altruistic 3PP should behave prosocially in different contexts⁴, assuming that the tendency to behave prosocially carries over different situations^{2,14–18}, and across time^{14,19,20}.

Developmental and physiological features have been found to influence cooperation^{21–25}, trustworthiness²⁶ and spontaneous generosity²⁷. Thus, it is likely that 3PP is also influenced by these features, which seem to link the genetic propensity to behave prosocially, established throughout the evolutionary process^{7,28}, to actual prosociality. Genes, jointly with the developmental environment, affect various biological features that in turn influence behavior. These features include hormonal levels, specific neural properties or preconfigured neural circuits, and even physiological features not directly related to behavior but with an indirect impact on it. Sex has been the only biological variable previously studied in relation to 3PP: it has been shown that men and women engage in 3PP with different intensity, with men punishing more harshly than women^{15,29,30}. On the other hand, meta-analyses show that there is no such sexual dimorphism in cooperative behavior^{31,32}, indicating that the relationship between cooperation and 3PP might be more complex than previously thought.

Among the physiology-related variables with a likely influence on 3PP, those related to cooperation are good candidates^{21–25} given the described association between altruistic 3PP and the tendency to cooperate^{33–35}. Interestingly, some of these variables are sexually dimorphic^{36,37}, or show differential impacts on cooperation by

¹Biology Department, Universidad Autónoma de Madrid. C/Darwin 2, 28029, Madrid, Spain. ²Laboratorio de Comportamiento Animal y Humano, Centro de Estudios Avanzados, Universidad de Playa Ancha, Valparaíso, Chile.

³Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Lisbon, Portugal. ⁴Department of Political Economy, King's College London, London, UK. Correspondence and requests for materials should be addressed to E.T. (email: enrique.turiegano@uam.es)

sex^{25,38,39}. Therefore, these features might be behind the described sexual differences in 3PP. Our aim in this work is to study for the first time how features related to individual development influence the tendency to engage in altruistic 3PP in the context of a social dilemma. In addition, we study whether these features can also explain the observed sex differences in altruistic 3PP.

Cooperative behavior is affected by the levels of sexual hormones to which an individual is exposed during development^{21,22,26,40–42}. Their influence is due to their organizational effects during this period, in which they conform and modify the anatomy and physiology of growing organisms^{43,44}, including the nervous system⁴⁵. They can affect adult behavior through this influence on the development of the nervous system. Sexual hormones have crucial organizational effects during two periods: prenatal and pubertal^{44,45}. The exposure to sexual hormones during these periods can be proxied with two widely-used morphological features: the second-to-fourth digit ratio (2D:4D) and facial masculinity/femininity^{46,47}.

2D:4D is influenced by the prenatal ratio of testosterone/estrogen activity³⁶. Lower ratios are indicative of higher exposure to testosterone during human prenatal development⁴⁸. Although this relationship has become controversial in recent years^{49,50}, there is still solid evidence of a link between prenatal hormone levels and 2D:4D^{36,51,52}. The association between 2D:4D and prosocial behaviors (e.g. generosity, rejection of injustice) has been described in diverse experimental settings^{21,27}. However, the link between 2D:4D and cooperative behavior is not fully understood yet. High values of 2D:4D are related to less cooperation in both sexes. Results are mixed though for high levels of cooperation, which have been associated to both medium and low 2D:4D values^{21,22}. These conflicting results might be accounted for by the role of 2D:4D as a modulator of the effect of other variables on behavior^{27,53,54}.

Facial dimorphism, usually referred to as facial masculinity/femininity in the literature, is a proxy for pubertal exposure to sexual hormones⁴⁷, with higher values of dimorphism indicating higher levels of estrogens and testosterone levels for women and men, respectively. Higher facial masculinity is negatively related to prosociality in men^{26,40–42} (depending on the social context²³), as it is associated in the laboratory with aggression⁴¹, unethical behavior²⁶, and cheating in negotiations⁴². No correlation has been found between facial femininity and prosociality in women²⁵.

Another developmental variable with an impact on cooperation is facial fluctuating asymmetry (FA). This measure is related to developmental instability⁵⁵, and it is frequently employed as a proxy for phenotypic quality⁵⁶, although with some limitations⁵⁷. More asymmetrical individuals are more cooperative and engage more often in prosocial behaviors^{22,25,39}, albeit these effects are stronger in men. FA is closely related to attractiveness^{58,59}, which in turn has been shown to be negatively related to prosociality in women^{25,60,61} but positively in men³⁸.

In this work, we explore whether physiology related variables influence altruistic 3PP in the context of a social dilemma, the Prisoner's Dilemma Game (PDG), and whether these variables can explain the sex differences observed in this form of punishment. We expect cooperative individuals, both men and women, to be more likely to engage in altruistic 3PP (*Hypothesis 1*). On the basis that altruistic 3PP is related to cooperation, we entertain three additional hypotheses. We expect higher values of 2D:4D to correlate with a lower tendency to engage in 3PP (*Hypothesis 2*), given that cooperativeness is in turn negatively related with 2D:4D. We also expect men with lower facial dimorphism to engage in 3PP more often, but not women (*Hypothesis 3*), as facial dimorphism has a differential effect on cooperation by sex. Finally, we expect more asymmetrical men and women to exert 3PP more often given that FA is positively correlated with cooperation, although this association should be stronger in men (*Hypothesis 4*). In addition, we test differences in the aforementioned variables between participants who cooperated and those who did not. The aim of our study is to broaden our knowledge of the variables that influence prosociality and the persistence of social norms in order to better understand the causes behind the differences across human social organization systems.

Results

Sexual differences in cooperation and 3PP. 67.32% of women and 62.20% of men in our sample chose to cooperate; sex differences were not significant ($X^2_1 = 2.32$; $p = 0.128$; $\varphi = 0.05$). Regarding expected behavior, 68.49% of women and 65.24% of men believed that their counterpart would cooperate. Sex differences in these expectations were not significant either ($X^2_1 = 0.96$; $p = 0.328$; $\varphi = 0.03$). Cooperative behavior correlated strongly with expected cooperation of the counterpart (women: $X^2_1 = 135.73$; $p < 0.001$; $\varphi = 0.52$; men: $X^2_1 = 69.66$; $p < 0.001$; $\varphi = 0.46$).

As observers, 20.35% of women decided to punish the defector who expected her counterpart to cooperate. This percentage was 26.22% in men. Sex differences in 3PP frequency were significant ($X^2_1 = 3.93$; $p = 0.048$; $\varphi = 0.14$).

Relationship between cooperation, beliefs about the other's behavior and 3PP. There was no relationship between cooperation and 3PP, either in men ($X^2_1 = 2.04$; $p = 0.154$; $\varphi = 0.08$), or women, although the latter was close to significance ($X^2_1 = 3.52$; $p = 0.061$; $\varphi = 0.09$). We found no relationship either between the counterpart's expected behavior and 3PP (women: $X^2_1 = 2.93$; $p = 0.087$; $\varphi = 0.08$; men: $X^2_1 = 0.67$; $p = 0.412$; $\varphi = 0.05$). However, 28.92% of male cooperators engaged in 3PP whereas only 18.02% of female cooperators did so ($X^2_1 = 8.84$; $p = 0.003$; $\varphi = 0.12$). There were no sex differences in the proportion of defecting participants who engaged in 3PP ($X^2_1 = 0.45$; $p = 0.503$; $\varphi = 0.04$). Values of these different proportions are included in Supplementary Table S1.

We classified participants in four categories according to their behavior in the two stages, PDG and 3PP (Fig. 1). There were sex differences in the distribution of subjects across these four categories ($X^2_3 = 11.45$; $p = 0.010$; $w = 0.12$).

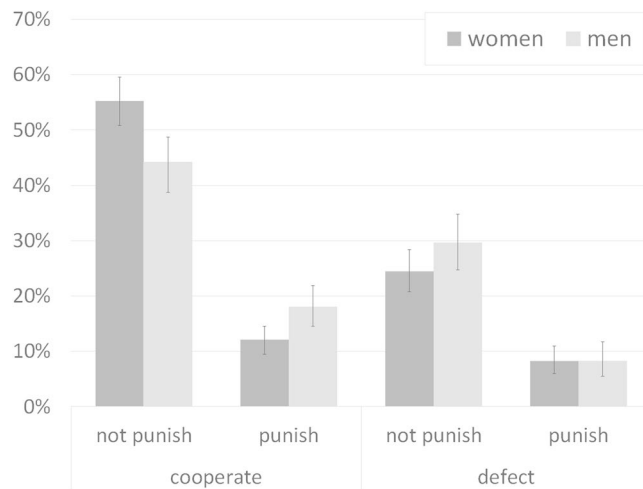


Figure 1. Proportion of male and female participants according to their 3PP and PDG choices. The bars represent the 95% confidence interval for the proportion.

Influence of biological variables on the tendency to punish third parties. We next analysed the influence of biological variables on 3PP by sex. Firstly, we checked for differences in men according to their punitive behavior. We found significant differences (Table 1) in facial dimorphism and self-perceived attractiveness (SPA). In women, we found no significant differences in any physiological variable between those who punished and those who did not.

We then pooled all observations and built a logistic model (the complete procedure followed is described in Supplementary Tables S2–S4). The final model included the variables sex, behavior in the PDG (PDGb), facial dimorphism and SPA, and their interactions with the sex dummy (Table 2). All variables showed a significant effect on punishing behavior mediated by sex. According to the estimates from the model, a man who cooperates is in average 1.27 times more likely to engage in 3PP compared with a defector, while a woman cooperator is in average 0.70 times more likely to punish than a woman who defected. On the other hand, the estimated risk ratios yield that men with a SPA one standard deviation over the mean are 1.26 times more likely to engage in 3PP, while women in the same situation are 0.85 times more likely to do so. In addition, men with one SD over the mean dimorphism are 0.76 times more likely to punish, while having one SD over the mean has a very weak effect on the likelihood of women engaging in 3PP (they are 1.06 times more likely).

Additional analysis: relationship between biological variables and cooperation. Given the differences in the relationship between cooperation and 3PP in men and women, we also checked for sex differences for cooperators and defectors separately in each of the variables considered (Table 1). There are no significant differences in expected behavior (Supplementary Table S5).

Discussion

One important result emerging from our analysis is that cooperation and 3PP are not associated equally in men and women. Although it is well described that men are harsher third-party punishers²⁹, no previous study has analysed the relationship between 3PP and sex controlling for cooperative behavior. A lack of correlation between cooperation and 3PP was previously described¹⁴, but without controlling for sex. In contrast, we find that cooperative men are more likely to punish defectors (risk ratio: 1.27), but cooperative women are not (risk ratio: 0.70). We observe this even after controlling for dimorphic variables which could explain this sex difference. In short, our *Hypothesis 1* is only supported for men. It is noteworthy that our design was aimed to reflect real altruism given that it entailed a real cost for the participants who exerted it^{13,62,63}.

This result suggests that the motivations of men and women to engage in 3PP could be different, a difference very relevant to the role of 3PP in supporting extended prosociality. It would be interesting to study whether this result is related to the previously described tendency of women to behave in a care-oriented and context-dependent way, in contrast to the rigid norm-oriented, context-independent behavior that is more prevalent in men^{64,65}. This differential way of evaluating third-parties could be driving the sex differences in 3PP described in the literature^{15,29,30}, a feature we also observe in our sample even after controlling for sexually dimorphic characteristics. It is interesting to point out that 3PP, in addition to promoting an increase in general prosociality⁶⁶, confers a gain in status and trustworthiness to the individual who performs it^{8,10}. This gain might be more appealing to men according to the stronger association they face between reproductive success and status⁶⁷. In contrast, second-party punishment does not award any of these advantages^{66,68}. Further analyses on the motivation to engage in 3PP in men and women controlling for the tendency to display care-oriented, normative, and status-seeking behaviors could help to clarify the sexual differences in the association between 3PP and cooperation and, consequently, on cooperative behavior within groups³².

In fact, the existence of sex differences in the motivations to engage in 3PP is also supported by the relationships between sexually dimorphic features and 3PP we observe. In line with *Hypothesis 3*, we find an inverse

	No Punish (N = 407)	Punish (N = 104)	Defect (N = 167)	Cooperate (N = 344)
WOMEN				
Dimorphism	4.54 (4.18–4.75)	4.78 (4.32–5.64)	4.70 (4.23–5.26)	4.54 (4.18–4.79)
	U = 19010, p = 0.109, d = -0.079			
FA	3.04 (2.94–3.03)	3.04 (2.91–3.11)	3.02 (2.91–3.04)	3.05 (2.94–3.06)
	U = 20905, p = 0.848, d = 0.001			
2D:4D	0.977 (0.973–0.979)	0.982 (0.975–0.989)	0.978 (0.973–0.981)	0.977 (0.974–0.981)
	U = 19158, p = 0.136, d = -0.182			
SPA	0.612 (0.589–0.639)	0.591 (0.570–0.616)	0.628 (0.610–0.647)	0.599 (0.576–0.619)
	U = 23260, p = 0.119, d = 0.148			
	No Punish (N = 242)	Punish (N = 86)	Defect (N = 124)	Cooperate (N = 204)
MEN				
Dimorphism	4.96 (4.57–5.32)	4.06 (3.30–4.85)	4.84 (4.29–5.43)	4.66 (4.23–5.08)
	U = 12019, p = 0.033, d = 0.294			
FA	2.86 (2.79–2.89)	2.77 (2.67–2.84)	2.82 (2.74–2.88)	2.85 (2.76–2.88)
	U = 11675, p = 0.093, d = 0.204			
2D:4D	0.965 (0.961–0.969)	0.965 (0.959–0.972)	0.967 (0.962–0.973)	0.964 (0.960–0.968)
	U = 10318, p = 0.907, d = -0.011			
SPA	0.592 (0.570–0.633)	0.657 (0.638–0.711)	0.587 (0.561–0.636)	0.623 (0.593–0.648)
	U = 8204, p = 0.004, d = -0.339			

Table 1. Mean values of variables according to participants' behavior in 3PP and PDG by sex. Data are presented as: mean (95% confidence interval).

association between facial dimorphism and 3PP in men (risk ratio: 0.76), which remains robust after controlling for related variables (PDGb, SPA). Let us highlight that the effect size of facial dimorphism on 3PP is similar to that of SPA, a self-evaluated variable. Higher facial masculinity is often related to lower prosociality in situations lacking social context^{26,40–42}, the same holds for other features related to pubertal testosterone^{60,69}. It thus makes sense that we observe higher facial dimorphism to be negatively related to 3PP in men since 3PP is considered to be a prosocial behavior. It is not infrequent to observe individuals who display features associated with a good physical condition to be less prosocial^{22,24,25,39}. This is often explained by their alleged greater ability to obtain resources by themselves^{17,70}.

The effect of SPA we describe supports the view of 3PP as a prosocial behavior: men who deem themselves as more attractive are more likely to engage in 3PP (risk ratio: 1.26). This agrees with previous studies showing that men who consider themselves attractive are more prosocial³⁸. We also observe this relationship in our PDG: men with higher SPA are indeed more likely to cooperate (d = 0.186; Table 1). The effect of SPA in men's behavior is consistent with the so-called "attractiveness halo effect"⁷¹, which states that more attractive people are treated better by others but are also expected to be more prosocial⁷².

In women, we expected no relationship between the degree of facial dimorphism and 3PP. This was based on the lack of a clear link between women's facial dimorphism and prosocial/antisocial behavior^{25,26,42}. The model in Table 2 shows a very weak relationship between facial dimorphism and 3PP in women (risk ratio: 1.06). In fact, this relationship is not significant at any standard confidence level before controlling for other variables (Table 1). SPA shows a negative effect on the likelihood to engage in 3PP when facial dimorphism and cooperation in the PDG are controlled for (risk ratio: 0.85). The standing literature offers conflicting results on the association between attractiveness and prosocial behavior in women; some studies show a negative association^{25,38,60}, whereas others show no relationship at all^{61,73}. Our study confirms these mixed results. We find no significant differences in SPA between women who punished and those who did not, but SPA shows a negative effect on 3PP in our final model. This difference is due to the fact that we control for PDG behavior: defection in the PDG is positively associated with 3PP (Table 2), and defecting women tend to see themselves as more attractive (d = 0.203; Table 1). Hence, once we control for behavior in the PDG, a negative relationship between SPA and 3PP emerges. It would be interesting to carry out further experiments on this complex relationship between SPA, prosociality and 3PP in women, with SPA being a variable of interest rather than just a control.

As mentioned, the sex differences we observe in the relationship between cooperation and 3PP and between facial dimorphism and 3PP suggest that men and women may have different underlying proximal or evolutionary motivations to engage in altruistic punishment. These differences could be worth exploring in the future; for instance, it would be interesting to understand better the differential impact of 3PP on fitness by sex or the influence of empathic mechanisms on this behavior. In any case, the lack of a common relationship between 3PP and cooperation across sexes indicates that 3PP might not be a mechanism enforcing cooperation within a community, but a form of exploitation. Further research must investigate how altruistic 3PP, a behavior exerted mostly and more intensively by half of the population, could support extended cooperation^{5,34} as some models suggest^{74,75} under different contexts^{76,77} and under different forms of punishment^{78,79}. In this sense, let us reiterate that our results have been obtained studying exclusively a university population within a Western culture. It would be important to confirm them in a wider range of ages, ethnicities and socio-cultural strata. It is also worth investigating to what extent non-kin prosociality is motivated by social norms or by affective reactions, and

Variables in the model	MODEL				VARIABLE				
	−2LL	Likelihood Ratio Test	Df	p	Variables	Coef	Wald	Df	p
<i>Sex, PDG b, Dimorphism, SPA, Sex x PDG b, Sex x Dimorphism, Sex x SPA</i>	871.83	11.88	3	0.007	Constant	−1.662	11.33	1	0.001
					Sex	0.742	1.15	1	0.284
					PDG b	−0.312	1.31	1	0.252
					Dimorphism	−0.109	6.52	1	0.011
					SPA	1.969	7.80	1	0.005
					Sex x PDG b	0.769	4.67	1	0.031
					Sex x Dimorphism	0.139	6.17	1	0.013
					Sex x SPA	−3.208	9.40	1	0.002

Table 2. Final logistic model for 3PP. The variables initially included were Dimorphism (facial dimorphism), SPA (self-perceived attractiveness), and the categorical variables sex and PDGb (behavior in the PDG). The reference values for the categorical variables are male for sex and cooperate for PDGb.

whether this relationship differs between sexes^{64,65}. It would also be interesting to delve deeper into the relationship between 3PP and other prosocial behaviors in order to establish whether 3PP exerted against an individual who harms others can be indeed considered prosocial¹⁴.

Some of the variables we expected to correlate with 3PP were not significant. We found no support in favour of *Hypothesis 2*: 2D:4D did not show any influence on behavior in either sex. Previous literature on 2D:4D and prosociality indicates that their association is non-linear^{21,22}, and that 2D:4D mediates the effect of other variables depending on the context^{27,53,54,80,81}. Future experiments varying the context in which decisions are made might uncover an effect of 2D:4D on 3PP. We do not find any evidence either of a direct effect of FA on 3PP (*Hypothesis 4*), another variable previously related to prosociality^{25,82}. This lack of a direct association indicates, once again, that the relationship between prosociality and 3PP is more complex than usually conjectured. Nonetheless, we observe a weak association between cooperation and FA in men ($d = 0.067$; Table 1) and between cooperation and punishment (Table 2). These results are consistent with the conjecture that altruistic punishment is a prosocial behavior in men but not in women.

We would like to emphasize at this point that although we chose our variables of interest considering that prosociality is stable over time and that individuals tend to behave prosocially in different contexts, this does not mean that prosociality must be considered as a trait. Stability across time and context could also result from prosociality being a group of correlated mechanisms leading to behaviors aimed at improving social welfare. In fact, our results suggest that prosociality is not a single trait, given that the set of variables related to the tendency to engage in altruistic 3PP and of those previously described to be related to prosociality do not overlap.

Before concluding, let us highlight a byproduct of our study. We carried our analysis in a very large sample, allowing us to test further the relationship between cooperation and some individual features. Our analysis confirm some of the results observed in moderate sample sizes, such as the relationship between FA and cooperation in men^{22,83}, and the positive (negative) correlation between SPA and cooperation in men (women)^{25,38}. We found no differences in the other variables considered (facial dimorphism, 2D:4D) between cooperators and non-cooperators in either men or women (Table 1).

Conclusions

Our findings provide new insights on the relationship between cooperation and altruistic punishment. There exist sex differences in the tendency to engage in 3PP, with cooperating men being more likely to punish. This contrasts with what we observe in women. We also explore how pubertal levels of sexual hormones influence altruistic punishment: as expected, there is an inverse association between facial dimorphism and 3PP in men, whereas the association is much weaker or non-existent in women. These findings suggest that the motivations for performing 3PP are different in men and women. We hope to have broadened the debate on the underlying causes of prosociality and their nuances.

Methods

Participants. 839 students from Universidad Autónoma de Madrid (Spain) participated in the experiment (511 women and 328 men) once only. Mean age (\pm SD) of women was 21.09 ± 2.57 , and mean age of men was 21.82 ± 2.32 . All participants included in this study classified themselves as Caucasian. Data from participants belonging to other ethnicities, marginally represented in this population, were collected but discarded because they introduced facial shape variations stronger than sexual dimorphism in the analysis.

Experimental session. Each session was carried out with 8–20 same-sex participants. Before the experiment took place, subjects were instructed about the procedures. They filled in a questionnaire eliciting their ethnicity and self-perceived attractiveness (SPA), measured in a slider scale. The experimental procedures received approval from the Universidad Autónoma de Madrid (UAM) Ethics Committee (approval number: CEI 73–1319). All the methods were performed in accordance with the relevant guidelines and regulations of the institution, and informed consent of all the participants was obtained.

	Sex	N	Mean \pm 95% CI	t	Df	p	d
Dimorphism	Women	511	4.59 \pm 0.27	-0.63	837	0.529	-0.045
	Men	328	4.72 \pm 0.34				
FA	Women	511	3.04 \pm 0.05	5.63	837	< 0.001	0.398
	Men	328	2.84 \pm 0.05				
2D:4D	Women	511	0.978 \pm 0.003	5.90	837	< 0.001	0.418
	Men	328	0.965 \pm 0.003				
SPA	Women	511	0.608 \pm 0.013	-0.10	837	0.918	-0.007
	Men	328	0.609 \pm 0.021				

Table 3. Descriptive statistics by sex. The variables included are Dimorphism (facial dimorphism), FA (Fluctuating asymmetry), 2D:4D (average of left and right second to fourth digit ratio), SPA (self-perceived attractiveness). 95% CI: 95% confidence interval for each mean.

The experiment was programmed using the z-Tree 3.2.10 software for economic experiments⁸⁴. In the first stage, participants played a prisoner's dilemma game (PDG), a strategic game widely used to measure cooperation^{12,19,22,24}. They had to decide whether to cooperate or not with an anonymous counterpart. This counterpart was an anonymous participant in a previous session. We informed participants that their decisions would be employed in future sessions as anonymous potential counterparts. We also asked our participants about the behavior they expected from their counterpart. Participants were not informed about the result of this interaction. In the second stage, participants played as observers. They were presented a PDG between two other individuals. Participants were asked what they would do if one of these individuals chose not to cooperate whilst declaring that he/she expected his/her counterpart to cooperate. The participant had to decide whether he/she would pay a cost to reduce the earnings of this defector but without knowing the actual outcome of the PDG they were observing. The individuals playing the PDG observed by our participants were subjects of future experimental sessions, so their punishment decisions had monetary consequences. We implemented the decisions of observers in a separate series of sessions whose data is not included in the present analysis.

Participants were paid according to their choices during the entire session, which included other experiments not analysed in the present study. In the PDG stage, payoffs were as follows: if they both cooperated, participants got 90 points; if none of them cooperated, they both got 30 points; if one of them cooperated and the other did not, the cooperator got 10 points, and the defector 160 points. In the 3PP stage, participants were assigned 80 points and were asked whether they would pay half of these points (40 points) to decrease by 120 points the total earnings of the uncooperative participant they were observing. Conversion from points to euros was 1€/100 points. The size of stakes was thus relatively small, a feature which could affect strategic behavior⁸⁵. We nonetheless employed these values as they had already been successfully implemented in similar studies^{22,25}. Let us emphasize that our design allows participants to engage in truly altruistic punishment^{13,62,63}. There was only one punishing stage, so there was no strategic motivation to exert 3PP; that is, there were no additional future interactions that participants could take into account when making their decisions. In addition, participants could not use 3PP to build up reputation because their choice was one-shot and was not made public. Their motivation could not be revenge either, since participants could not punish individuals they had interacted with at any moment. Punishments could not be affected either by distributional concerns such as spite or inequality aversion, as participants were not informed of the outcome of the PDG they were observing. Finally, let us reiterate that punishment decisions were costly to our participants and had an actual effect on other participants; they were not hypothetical and had real consequences.

Measurements. We obtained hand-scans from each participant with a CanoScan LiDE 200 scanner as well as two frontal facial photographs in standardized conditions. Lengths of the second and fourth digits were measured from the hand scans as it is customary in the literature^{21,86,87}. We computed their ratios, averaged the two resulting numbers, and included them in the analysis. In order to compute the facial FA and facial dimorphism measures, two independent observers placed 39 predefined landmarks in the two photographs of each participant with the TPS 2.16 software (by F.J. Rohlf; obtained from <http://life.bio.sunysb.edu/ee/rohlf/software.html>). Morphometric analyses were carried out with MorphoJ 1.04a (by C.P. Klingenberg; obtained from http://www.flywings.org.uk/morphoj_page.htm). The Mahalanobis FA score was computed, using the methods described in the literature²². Dimorphism scores were obtained by a discriminant score as described in previous studies³⁷; women's scores were converted to positive values to obtain an absolute value of facial dimorphism. Higher scores indicate higher facial shape differences from the other sex. Descriptive statistics for men and women are shown in Table 3.

Statistical analyses. Statistical analyses were carried out with SPSS 15.0. We checked the normality of all variables and of the residuals in all models.

We analysed the relationship between biological variables and behavior by means of t-tests. The effect of squared 2D:4D was tested simultaneously with that of 2D:4D²² in a logistic regression, but the effect was not significant (Supplementary Table S6).

Models were built following standard recommendations^{88,89}, namely testing each variable independently and introducing all the significant ones and those with a strong theoretical support in one model for men and women separately. We included the choice and the expected behavior of the counterpart in the first stage PDG,

considering the behavior displayed as a measure of cooperativeness. Next, we tested those variables in a single model including their interaction with sex, searching for the simpler significant model. Finally, we checked the interactions between the remaining variables.

As it is recommended for logistic models⁹⁰, we computed effect sizes for all variables in the form of risk ratios, i.e. the ratio between the probability of the event before and after applying a change in the selected variable. For continuous variables, the change consists of a one standard deviation increase from the mean of the variable. For dichotomous variables, the change is just from 0 to 1. A comparison between the effect size after and before their inclusion in the model are showed in Supplementary Fig. S1.

Data Availability

Data from this study is available as Supplementary Material.

References

- Fehr, E. & Gächter, S. Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
- Henrich, J. *et al.* “Economic man” in cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behav. Brain Sci.* **28**, 795–855 (2005).
- Henrich, J., Heine, S. J. & Norenzayan, A. The weirdest people in the world? *Behav. Brain Sci.* **33**, 61–135 (2010).
- Gächter, S. & Herrmann, B. Reciprocity, culture and human cooperation: Previous insights and a new cross-cultural experiment. *Philos. T. Roy. Soc. B* **364**, 791–806 (2009).
- Boyd, R., Gintis, H., Bowles, S. & Richerson, P. The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. USA* **100**, 3531–3535 (2003).
- Henrich, J. *et al.* Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).
- West, S. A., El Mouden, C. & Gardner, A. Sixteen common misconceptions about the evolution of cooperation in humans. *Evol. Hum. Behav.* **32**, 231–262 (2011).
- Nelissen, R. M. A. The price you pay: Cost-dependent reputation effects of altruistic punishment. *Evol. Hum. Behav.* **29**, 242–248 (2008).
- Kurzban, R., DeScioli, P. & O’Brien, E. Audience effects on moralistic punishment. *Evol. Hum. Behav.* **28**, 75–84 (2007).
- Barclay, P. Reputational benefits for altruistic punishment. *Evol. Hum. Behav.* **27**, 325–344 (2006).
- Raihani, N. J. & Bshary, R. Third-party punishers are rewarded, but third-party helpers even more so. *Evolution* **69**, 993–1003 (2015).
- Rand, D. G. & Nowak, M. A. Human cooperation. *Trends Cogn. Sci.* **17**, 413–425 (2013).
- Kurzban, R., Burton-Chellew, M. N. & West, S. A. The Evolution of Altruism in Humans. *Annu. Rev. Psychol.* **66**, 575–599 (2015).
- Peysakhovich, A., Nowak, M. A. & Rand, D. G. Humans display a ‘cooperative phenotype’ that is domain general and temporally stable. *Nat. Commun.* **5**, 4939, <https://doi.org/10.1038/ncomms5939> (2014).
- Eckel, C. & Grossman, P. The relative price of fairness: Gender differences in a punishment game. *J. Econ. Behav. Organ.* **30**, 143–158 (1996).
- Qin, X., Shen, J. & Meng, X. Group-based trust, trustworthiness and voluntary cooperation: Evidence from experimental and survey data in China. *J. Socio-Econ.* **40**, 356–363 (2011).
- Sanchez-Pages, S. & Turiegano, E. Two studies on the interplay between social preferences and individual biological features. *Behaviour* **150**, 713–735 (2013).
- Yamagishi, T., Tanida, S., Mashima, R., Shimoma, E. & Kanazawa, S. You can judge a book by its cover - Evidence that cheaters may look different from cooperators. *Evol. Hum. Behav.* **24**, 290–301 (2003).
- Mao, A., Dworkin, L., Suri, S. & Watts, D. J. Resilient cooperators stabilize long-run cooperation in the finitely repeated Prisoner’s Dilemma. *Nat. Commun.* **8**, 13800, <https://doi.org/10.1038/ncomms13800> (2017).
- Carlsson, F., Johansson-Stenman, O. & Nam, P. K. Social preferences are stable over long periods of time. *J. Public Econ.* **117**, 104–114 (2014).
- Millet, K. & Dewitte, S. Second to fourth digit ratio and cooperative behavior. *Biol. Psychol.* **71**, 111–115 (2006).
- Sanchez-Pages, S. & Turiegano, E. Testosterone, facial symmetry and cooperation in the prisoners’ dilemma. *Physiol. Behav.* **99**, 355–361 (2010).
- Stirrat, M. & Perrett, D. I. Face structure predicts cooperation: Men with wider faces are more generous to their in-group when out-group competition is salient. *Psychol. Sci.* **23**, 718–722 (2012).
- Shinada, M. & Yamagishi, T. Physical attractiveness and cooperation in a prisoner’s dilemma game. *Evol. Hum. Behav.* **35**, 451–455 (2014).
- Munoz-Reyes, J. A., Pita, M., Arjona, M., Sanchez-Pages, S. & Turiegano, E. Who is the fairest of them all? The independent effect of attractive features and self-perceived attractiveness on cooperation among women. *Evol. Hum. Behav.* **35**, 118–125 (2014).
- Stirrat, M. & Perrett, D. I. valid facial cues to cooperation and trust: Male facial width and trustworthiness. *Psychol. Sci.* **21**, 349–354 (2010).
- Millet, K. & Dewitte, S. The presence of aggression cues inverts the relation between digit ratio (2D:4D) and prosocial behaviour in a dictator game. *Br. J. Psychol.* **100**, 151–162 (2009).
- Ebstein, R. P., Israel, S., Chew, S. H., Zhong, S. & Knafo, A. Genetics of human social behavior. *Neuron* **65**, 831–844 (2010).
- Kromer, E. & Bahçekapili, H. G. The influence of cooperative environment and gender on economic decisions in a third party punishment game. *Wpcpg 2010* **5**, 250–254 (2010).
- Li, J. & Zhou, X. Sex, attractiveness, and third-party punishment in fairness consideration. *PLoS One* **9**, e94004, <https://doi.org/10.1371/journal.pone.0094004> (2014).
- Balliet, D., Parks, C. & Joireman, J. Social value orientation and cooperation in social dilemmas: a meta-analysis. *Group Processes Interg.* **12**, 533–547 (2009).
- Balliet, D., Li, N. P., Macfarlan, S. J. & Van Vugt, M. Sex differences in cooperation: A meta-analytic review of social dilemmas. *Psychol. Bull.* **137**, 881–909 (2011).
- Fehr, E., Fischbacher, U. & Gächter, S. Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum. Nature* **13**, 1–25 (2002).
- Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425**, 785–791 (2003).
- Bowles, S. & Gintis, H. The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theor. Popul. Biol.* **65**, 17–28 (2004).
- Zheng, Z. & Cohn, M. J. Developmental basis of sexually dimorphic digit ratios. *Proc. Natl. Acad. Sci. USA* **108**, 16289–16294 (2011).
- Sanchez-Pages, S., Rodríguez-Ruiz, C. & Turiegano, E. Facial masculinity: How the choice of measurement method enables to detect its influence on behaviour. *PLoS One* **9**, e112157, <https://doi.org/10.1371/journal.pone.0112157> (2014).
- Mulford, M., Orbell, J., Shatto, C. & Stockard, J. Physical attractiveness, opportunity, and success in everyday exchange. *Am. J. Sociol.* **103**, 1565–1592 (1998).
- Zaatari, D. & Trivers, R. Fluctuating asymmetry and behavior in the ultimatum game in Jamaica. *Evol. Hum. Behav.* **28**, 223–227 (2007).

40. Haselhuhn, M. P. & Wong, E. M. Bad to the bone: Facial structure predicts unethical behaviour. *Proc. R. Soc. B-Biol. Sci.* **279**, 571–576 (2012).
41. Goetz, S. M. M. *et al.* Social status moderates the relationship between facial structure and aggression. *Psychol. Sci.* **24**, 2329–2334 (2013).
42. Geniole, S. N., Keyes, A. E., Carre, J. M. & McCormick, C. M. Fearless dominance mediates the relationship between the facial width-to-height ratio and willingness to cheat. *Pers. Individ. Differ.* **57**, 59–64 (2014).
43. Sisk, C. & Zehr, J. Pubertal hormones organize the adolescent brain and behavior. *Front. Neuroendocrinol.* **26**, 163–174 (2005).
44. Schulz, K. M., Molenda-Figueira, H. A. & Sisk, C. L. Back to the future: The organizational-activational hypothesis adapted to puberty and adolescence. *Horm. Behav.* **55**, 597–604 (2009).
45. Berenbaum, S. A. & Beltz, A. M. Sexual differentiation of human behavior: Effects of prenatal and pubertal organizational hormones. *Front. Neuroendocrinol.* **32**, 183–200 (2011).
46. Manning, J., Bundred, P., Newton, D. & Flanagan, B. The second to fourth digit ratio and variation in the androgen receptor gene. *Evol. Hum. Behav.* **24**, 399–405 (2003).
47. Welker, K. M., Bird, B. & Arnoocky, S. Commentary: Facial Width-to-Height Ratio (fWHR) is not associated with adolescent testosterone levels. *Front. Psychol.* **7**, 1745, <https://doi.org/10.3389/fpsyg.2016.01745> (2016).
48. Manning, J., Bundred, P. & Flanagan, B. The ratio of 2nd to 4th digit length: A proxy for transactivation activity of the androgen receptor gene? *Med. Hypotheses* **59**, 334–336 (2002).
49. Hickey, M. *et al.* Maternal and umbilical cord androgen concentrations do not predict digit ratio (2D:4D) in girls: a prospective cohort study. *Psychoneuroendocrinol.* **35**, 1235–1244 (2010).
50. Hollier, L. P. *et al.* Adult digit ratio (2D:4D) is not related to umbilical cord androgen or estrogen concentrations, their ratios or net bioactivity. *Early Hum. Dev.* **91**, 111–117 (2015).
51. Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R. & Manning, J. 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Hum. Dev.* **77**, 23–28 (2004).
52. Ventura, T., Gomes, M., Pita, A., Neto, M. & Taylor, A. Digit ratio (2D:4D) in newborns: influences of prenatal testosterone and maternal environment. *Early Hum. Dev.* **89**, 107–112 (2013).
53. van Honk, J., Montoya, E. R., Bos, P. A., van Vugt, M. & Terburg, D. New evidence on testosterone and cooperation. *Nature* **485**, E4–E5 (2012).
54. Van den Bergh, B. & Dewitte, S. Digit ratio (2D:4D) moderates the impact of sexual cues on men's decisions in ultimatum games. *Proc. R. Soc. B-Biol. Sci.* **273**, 2091–2095 (2006).
55. Van Dongen, S. & Gangestad, S. W. Human fluctuating asymmetry in relation to health and quality: a meta-analysis. *Evol. Hum. Behav.* **32**, 380–398 (2011).
56. Thornhill, R. & Moller, A. Developmental stability, disease and medicine. *Biol. Rev.* **72**, 497–548 (1997).
57. Graham, J. H., Raz, S., Hel-Or, H. & Nevo, E. Fluctuating asymmetry: methods, theory, and applications. *Symmetry* **2**, 466–540 (2010).
58. Rhodes, G. The evolutionary psychology of facial beauty. *Annu. Rev. Psychol.* **57**, 199–226 (2006).
59. Kościński, K. Facial attractiveness: General patterns of facial preferences. *Anthropol. Rev.* **70**, 45–79 (2007).
60. Price, M. E., Kang, J., Dunn, J. & Hopkins, S. Muscularity and attractiveness as predictors of human egalitarianism. *Pers. Individ. Differ.* **50**, 636–640 (2011).
61. Bhogal, M. S., Galbraith, N. & Manktelow, K. physical attractiveness, altruism and cooperation in an ultimatum game. *Curr. Psychol.* **36**, 549–555 (2017).
62. Pedersen, E. J., Kurzban, R. & McCullough, M. E. Do humans really punish altruistically? A closer look. *Proc. R. Soc. B-Biol. Sci.* **280**, 20122723, <https://doi.org/10.1098/rspb.2012.2723> (2013).
63. Leibbrandt, A. & Lopez-Perez, R. An exploration of third and second party punishment in ten simple games. *J. Econ. Behav. Organ.* **84**, 753–766 (2012).
64. Jaffee, S. & Hyde, J. Gender differences in moral orientation: A meta-analysis. *Psychol. Bull.* **126**, 703–726 (2000).
65. Friesdorf, R., Conway, P. & Gawronski, B. Gender differences in responses to moral dilemmas: a process dissociation analysis. *Person. Soc. Psychol. Bull.* **41**, 696–713 (2015).
66. Zhou, Y., Jiao, P. & Zhang, Q. Second-party and third-party punishment in a public goods experiment. *Appl. Econ. Lett.* **24**, 54–57 (2017).
67. Puts, D. A. Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* **31**, 157–175 (2010).
68. Przepiorka, W. & Liebe, U. Generosity is a sign of trustworthiness—the punishment of selfishness is not. *Evol. Hum. Behav.* **37**, 255–262 (2016).
69. Price, M. E., Sheehy-Skeffington, J., Sidnaius, J. & Pound, N. Is sociopolitical egalitarianism related to bodily and facial formidability in men? *Evol. Hum. Behav.* **38**, 626–634 (2017).
70. Apicella, C. L. *et al.* Testosterone and financial risk preferences. *Evol. Hum. Behav.* **29**, 384–390 (2008).
71. Feingold, A. Good-looking people are not what we think. *Psychol. Bull.* **111**, 304–341 (1992).
72. Wilson, R. K. & Eckel, C. C. Judging a book by its cover: Beauty and expectations in the trust game. *Polit. Res. Quart.* **59**, 189–202 (2006).
73. Takahashi, C., Yamagishi, T., Tanida, S., Kiyonari, T. & Kanazawa, S. Attractiveness and cooperation in social exchange. *Evol. Psychol.* **4**, 315–329 (2006).
74. Chen, X., Szolnoki, A. & Perc, M. Competition and cooperation among different punishing strategies in the spatial public goods game. *Phys. Rev. E* **92**, 012819, <https://doi.org/10.1103/PhysRevE.92.012819> (2015).
75. Yang, H. & Chen, X. Promoting cooperation by punishing minority. *Appl. Math. Comput.* **316**, 460–466 (2018).
76. Chen, X. & Szolnoki, A. Punishment and inspection for governing the commons in a feedback-evolving game. *PLoS Comput. Biol.* **14**, e1006347, <https://doi.org/10.1371/journal.pcbi.1006347> (2018).
77. Wang, Q., He, N. & Chen, X. Replicator dynamics for public goods game with resource allocation in large populations. *Appl. Math. Comput.* **328**, 162–170 (2018).
78. Chen, X., Szolnoki, A. & Perc, M. Probabilistic sharing solves the problem of costly punishment. *New J. Phys.* **16**, 083016, <https://doi.org/10.1088/1367-2630/16/8/083016> (2014).
79. Liu, L., Chen, X. & Szolnoki, A. Competitions between prosocial exclusions and punishments in finite populations. *Sci. Rep.* **7**, 46634, <https://doi.org/10.1038/srep46634> (2017).
80. Millet, K. An interactionist perspective on the relation between 2D:4D and behavior: An overview of (moderated) relationships between 2D:4D and economic decision making. *Pers. Individ. Differ.* **51**, 397–401 (2011).
81. Montoya, E. R. *et al.* Testosterone administration modulates moral judgments depending on second-to-fourth digit ratio. *Psychoneuroendocrinol.* **38**, 1362–1369 (2013).
82. Munoz-Reyes, J. A., Gil-Burmann, C., Fink, B. & Turiegano, E. Facial asymmetry and aggression in Spanish adolescents. *Pers. Individ. Differ.* **53**, 857–861 (2012).
83. Munoz-Reyes, J. A., Gil-Burmann, C. & Turiegano, E. Digit Ratio 2D:4D, facial masculinization and aggressiveness in Spanish adolescents. *Estud. Psicol.* **35**, 319–340 (2014).
84. Fischbacher, U. z-Tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* **10**, 171–178 (2007).

85. Zhang, Y., Chen, X., Liu, A. & Sun, C. The effect of the stake size on the evolution of fairness. *Appl. Math. Comput.* **321**, 641–653 (2018).
86. Apicella, C. L. *et al.* Androgens and competitiveness in men. *J. Neurosci. Psychol. E.* **4**, 54–62 (2011).
87. Pearson, M. & Schipper, B. C. The visible hand: finger ratio (2D:4D) and competitive bidding. *Exp. Econ.* **15**, 510–529 (2012).
88. Harrell, F. E. In *Regression modeling strategies* (Springer-Verlag, Berlin, 2001).
89. Kleinbaum, D. G. & Klein, M. In *Logistic regression: A self-learning text* (Springer, New York, 2010).
90. Durlak, J. A. How to select, calculate, and interpret effect sizes. *J. Pediatr. Psychol.* **34**, 917–928 (2009).

Acknowledgements

Funding from the Spanish Ministry for Science and Innovation (grant ECO2015-66281-P) and from Chilean FONDECYT (Grant No. 1170513) are gratefully acknowledged.

Author Contributions

C.R.R. was involved in programming z-Tree stages, recruiting participants, carrying out the experimental sessions, analysing the data and writing the manuscript. J.A., M.R. was involved in programming z-Tree stages, carrying out experimental sessions and revising the manuscript. M.I.J. was involved in carrying out the experimental sessions and revising the manuscript. S.S.P. was involved in experimental design, programming z-Tree stages, carrying out experimental sessions and writing the manuscript. E.T. was involved in experimental design, carrying out experimental sessions, analysing the data and writing the manuscript. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-40909-8>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019