Research Article

Psychological Science
2018, Vol. 29(6) 868–876

DOI: 10.1177/0956797617752642
www.psychologicalscience.org/PS
SAGE

Why Hate the Good Guy? Antisocial Punishment of High Cooperators Is Greater When People Compete To Be Chosen

Aleta Pleasant and Pat Barclay
Department of Psychology, University of Guelph

Abstract

When choosing social partners, people prefer good cooperators (all else being equal). Given this preference, people wishing to be chosen can either increase their own cooperation to become more desirable or suppress others' cooperation to make them less desirable. Previous research shows that very cooperative people sometimes get punished ("antisocial punishment") or criticized ("do-gooder derogation") in many cultures. Here, we used a public-goods game with punishment to test whether antisocial punishment is used as a means of competing to be chosen by suppressing others' cooperation. As predicted, there was more antisocial punishment when participants were competing to be chosen for a subsequent cooperative task (a trust game) than without a subsequent task. This difference in antisocial punishment cannot be explained by differences in contributions, moralistic punishment, or confusion. This suggests that antisocial punishment is a social strategy that low cooperators use to avoid looking bad when high cooperators escalate cooperation.

Keywords

public-goods games, biological markets, partner choice, competitive altruism, do-gooder derogation, open data, open materials

Received 6/15/17; Revision accepted 11/13/17

> Le mieux est l'ennemi du bien (The best is the enemy of the good).
>
> —Voltaire (1865/1772, p. 541)

Cooperation is crucial to human success but is often personally costly (Barclay & Van Vugt, 2015). Noncooperators are often verbally reprimanded, excluded from cooperative interactions, or ostracized completely (e.g., Delton, Nemirow, Robertson, Cimino, & Cosmides, 2013; Feinberg, Willer, & Schultz, 2014), and this *moralistic punishment* can foster group cooperation because it motivates noncooperators to cooperate (see the meta-analysis by Balliet, Mulder, & Van Lange, 2011). However, highly cooperative individuals usually also get punished (Cinyabuguma, Page, & Putterman, 2006; Kuběna, Houdek, Lindová, Příplatová, & Flegr, 2014). This *antisocial punishment* of cooperators is found in every society tested so far (Herrmann, Thöni, & Gächter, 2008), as well as in field experiments (Barr, 2001). In experiments, approximately 20% of all punishment is used antisocially and tends to occur when the lowest cooperators target the highest cooperators (Cinyabuguma et al., 2006; Kuběna et al., 2014) or remove them from the group (Parks & Stone, 2010). Similar phenomena are found by other researchers: Social psychologists find *do-gooder derogation*, whereby people who help others get criticized or ridiculed for their efforts (Minson & Monin, 2012; Monin, 2007), anthropologists find that good hunters sometimes get criticized when they catch large shareable game (Boehm, 1999), and workers in organizations are sometimes criticized for working harder than others. Why?

Antisocial punishment and do-gooder derogation are puzzling because they destabilize group cooperation, reduce the beneficial effects of moralistic punishment,

Corresponding Author:
Pat Barclay, Department of Psychology, University of Guelph, 50 Stone Rd. East, Guelph, Ontario, N1G 2W1, Canada
E-mail: barclayp@uoguelph.ca

and, thus, lower payoffs for the entire group including the punisher (Herrmann et al., 2008; Rand & Nowak, 2011). The function of antisocial punishment is currently unknown. Some antisocial punishment might be retaliation by low cooperators against those whom they think punished them or might do so in the future (Herrmann et al., 2008; Sylwester, Herrmann, & Bryson, 2013; Van Dijk, Molenmaker, & Kwaadsteniet, 2015). Theories of normative conformity suggest that norm violators are punished simply for deviating from group norms, whether those deviations involve cooperating too little or too much (Henrich, 2004; Van Dijk et al., 2015). Antisocial punishment might also be a spiteful way of reducing others' payoff in order to have a higher payoff than them (Fliessbach et al., 2007); such spite is especially useful in areas with high economic inequality or competition over scarce resources, which is where antisocial punishment tends to be high (Fehr, Hoff, & Kshetramade, 2008; Prediger, Vollan, & Herrmann, 2014; Sylwester et al., 2013).

Barclay (2013) proposed a different hypothesis about antisocial punishment and do-gooder derogation, which is based on biological-markets theory. In biological-markets theory (Noë & Hammerstein, 1994, 1995), organisms can choose partners for cooperative interactions. The best partners are those who are most able, willing, and available to provide benefits in cooperative interactions (Barclay, 2013, 2016). This usually leads to organisms outbidding each other in order to be chosen as partners (Noë & Hammerstein, 1994), for example, by engaging in competitive helping or competitive altruism to attract social partners (Barclay, 2011, 2013; Roberts, 1998). Evidence of competitive helping has been demonstrated experimentally in humans (Barclay & Willer, 2007; Sylwester & Roberts, 2010) and in field studies of online donations (Raihani & Smith, 2015). However, outbidding competitors is only one way to compete: Organisms can also make their competitors look bad or prevent them from cooperating in order to look better by comparison (Barclay, 2013, 2016). Antisocial punishment and do-gooder derogation are simply two ways to accomplish this goal.

Attacks on high-value individuals occur in other domains. For example, people compete over romantic partners both by making themselves look good and by derogating competitors (Buss & Dedden, 1990). Barclay (2013) proposed that this tendency to derogate or attack competitors generalizes to other types of cooperative partnerships. Market value—just like mate value—is a relative construct: If your competitor looks very good, you look bad by comparison, and vice versa. The best cooperators make the moderate cooperators look bad—hence "the best is the enemy of the good." The low and moderate cooperators thus have an incentive in attacking or preventing the cooperation or morality of the best cooperators. This can occur in group cooperation, in philanthropy, within organizations (e.g., "you're working too hard"), or anytime one person has the potential to be perceived as morally superior to others (e.g., criticism of vegetarians, activists, or religious people). This is not limited to humans: Zahavi and Zahavi (1997) suggested that members of cooperatively breeding bird species will prevent group members from helping at the nest or feeding the chicks to prevent them from gaining social status for doing so. Thus, antisocial punishment may be an attempt to stop high cooperators from looking too good, to force them to cooperate less, and, by extension, to stop the antisocial punisher from looking selfish in comparison. We do not suggest that people do this consciously, any more than birds do. Instead, cues for partner choice will upregulate whatever psychological mechanisms cause competitive helping and antisocial punishment (for the distinction between psychological mechanisms and ultimate function, especially as they pertain to cooperation and punishment, see Barclay, 2012; Scott-Phillips, Dickins, & West, 2011; Tinbergen, 1968).

If Barclay's (2013) hypothesis is true, and if antisocial punishment is a way of competing over cooperative partnerships, then we would predict more antisocial punishment when people are explicitly competing over cooperative partnerships. We tested this prediction using a common laboratory measure of cooperation and punishment, in which participants were given the options to contribute to a public good and to punish others. We had two conditions. In the biological-markets condition, a third-party individual, called the "observer," saw the overall average contributions of each player and then decided whom he or she would like to play a subsequent economic game with (a trust game). The presence of the observer created competition to look more cooperative than others in the group in order to be chosen. The second condition was a control condition, which was identical to the biological-markets condition, except that no observer was present. We predicted more antisocial punishment in the biological-markets condition because low contributors would have more need to "bring down" the cooperation of the high contributors.

Method

Participants, earnings, and confidentiality

We recruited 75 females and 42 males (mean age = 19.6 years, *SD* = 2.0) from the University of Guelph psychology participant pool. Participants received psychology course credit plus lab dollars (L$), which were converted after the experiment to Canadian dollars (CAN$) at the preannounced rate of 10:1; earnings ranged from CAN$0 to CAN$15.25 (*M* = CAN$7.01, *SD* = CAN$2.59). Participants sat at computers with dividers to prevent visual contact with others, and the experiment was run using z-Tree software (Fischbacher, 2007). Participants received code numbers to preserve the anonymity of their decisions; these codes did not change across rounds. The experimenter knew participants' earnings but not the decisions that led to those earnings. All participants completed a comprehension test before making any decisions; the program clarified the rules to participants who got any particular question wrong. These methods were approved by the University of Guelph Research Ethics Board.

Procedure

Public-goods game (PGG) with punishment. Participants played five rounds of a PGG with punishment in groups of four people. In each round, each group member received L$10 and could contribute any number of these to a public good and keep the rest for him- or herself. The experimenter doubled all contributions and redistributed them evenly among all group members (contributors and noncontributors alike). This ensured that the group did best if everyone contributed, but each individual did best for him- or herself by keeping his or her own money and hoping that others contributed—individuals got only an L$0.50 private share of each L$1.00 they contributed.

After finding out what everyone contributed in that round, participants had the option to punish others by reducing their earnings: Every L$1 spent on punishment reduced the target's earnings by L$3. This money was not gained; both players lost money. Instead of calling it "punishment," we used the more neutral phrase "reducing others' earnings" to reduce framing effects. Participants could spend as much money on punishment as they had remaining that round, and they could punish any other group member. Participants were not told who punished them, only the total that they had been punished. After participants received the punishment, the round ended, and participants began the next round with a new L$10. Earnings were cumulative across all five rounds. Participants were not told the number of rounds.

Experimental conditions. The control condition was the standard PGG described above. The markets condition was the same as the control condition, except that there was a fifth participant in each session (the observer) who did not play the PGG but would later choose one of the other participants to play a trust game with, as follows. When the PGG finished, the observer was told how much each group member had contributed to the public good (using code numbers to preserve anonymity); observers were not told how much each person punished. The observer received L$50 and could entrust any amount to one group member that he or she chose (the "trustee"). The experimenter tripled any money entrusted, and the trustee could then decide how much to return to the observer (returns were not tripled). To collect trustee data from all PGG participants, we first asked what percentage of the entrusted money they would return in a hypothetical trust game. We then used the strategy method to elicit returns for the real trust game: All PGG participants were told how much was actually entrusted and were asked how much they wanted to return; if they were actually chosen, then their response would be implemented. Hypothetical returns and strategy-method returns were highly correlated after controlling for amount entrusted, partial $r(43) = .75$, $p < .001$, though data were missing from 6 participants for the latter because of a computer error.

Assignment to condition was not random but was unrelated to participant characteristics: We ran the control condition when 4 participants arrived on time to the session ($n = 52$ participants, 13 groups) and the markets condition when 5 or more participants arrived on time ($n = 65$ participants, 13 groups). We used this assignment method because participants occasionally do not show up, and this assignment method allowed us to maximize participants and minimize canceled sessions because of no-shows. Because participants signed up as individuals and did not know the experimental details until they arrived, this assignment method could not create a systematic difference in participant traits. Participants' responses to postexperimental surveys (see below) confirmed that participants were equivalent in the two conditions. Our sample size was determined by participant availability: We decided in advance to run as many sessions as possible in one semester, provided that we had at least 10 groups in each condition. A post hoc power analysis showed that our total of 26 sessions would have more than 80% power to find a large effect size (Cohen's $d = 0.8$).

Postexperimental surveys. While the experimenter calculated earnings from the PGG (and trust game in the markets condition), participants completed a short questionnaire. This included their age, gender, the extent to which they felt in competition with other participants (manipulation check on a Likert scale from 0 to 3), a hypothetical decision about whether they would prefer to play a new PGG for L$10 with a new group or receive a fixed L$12 for not participating ("loner question"), and six items from the Cooperative and Competitive Personality Scale (CCPS; Lu, Au, Jiang, Xie, & Yam, 2013; Xie, Yu, Chen, & Chen, 2006). The items from the CCPS included three questions from the cooperative subscale (e.g., "I enjoy working with other team members to achieve common success") and three questions from the competitive subscale (e.g., "Even in a group working towards a common goal, I still want to outperform others").

Results

Statistical analysis

Within a group of 4 in the PGG, each person's behavior affected everyone else's. To resolve this nonindependence of data, we treated each PGG group as $n = 1$ (e.g., Barclay, 2004; Nikiforakis & Normann, 2008). Punishment was coded as either moralistic or antisocial. Moralistic punishment was any instance of punishment targeting someone who contributed less than the group's average contribution that round, whereas antisocial punishment was any instance of punishment targeting someone who contributed more than (or equal to) the group's average contribution that round. To test the effects of market competition on punishment, we used a 2 × 5 mixed repeated measures general linear model (IBM SPSS 24) with the two experimental conditions (control, markets) as the between-subjects factor and the five rounds as a within-subjects factor. All violations of sphericity were corrected with the conservative Greenhouse-Geisser correction. Equal-variance *t* tests were used unless Levene's test indicated unequal variances.

Primary analysis: effect of partner choice on punishment and contributions

As predicted, antisocial punishment was higher in the markets condition than in the control condition, $F(1, 24) = 14.76$, $p = .001$, $\eta_p^2 = .38$ (Fig. 1a). Antisocial punishment did not change across rounds, nor was there a Round × Condition interaction (both *F*s < 1, both η_p^2s < .05). To test the robustness of the statistics, we noted that the average antisocial punishment per round was higher in the markets condition (*M* = L$3.05, *SE* = L$0.61) than in the control condition (*M* = L$0.63, *SE* = L$0.16), using an unequal-variance *t* test, $t(13.76) = 3.84$, $p = .002$, $d = 1.51$, 95% CI for the mean difference = [L$1.04, L$3.77], 95% CI for d = [0.59, 2.32]. The markets and control conditions differed in antisocial punishment right from the first round (*M* = L$3.46, *SE* = L$0.80 vs. *M* = L$0.46, *SE* = L$0.24, respectively); unequal-variance $t(14.21) = 3.60$, $p = .003$, $d = 1.41$, 95% CI for the mean difference = [L$1.21, L$4.78], 95% CI for d = [0.51, 2.22].

Our main result supports the hypothesis that participants use antisocial punishment to suppress others' generosity (and make themselves more competitive by

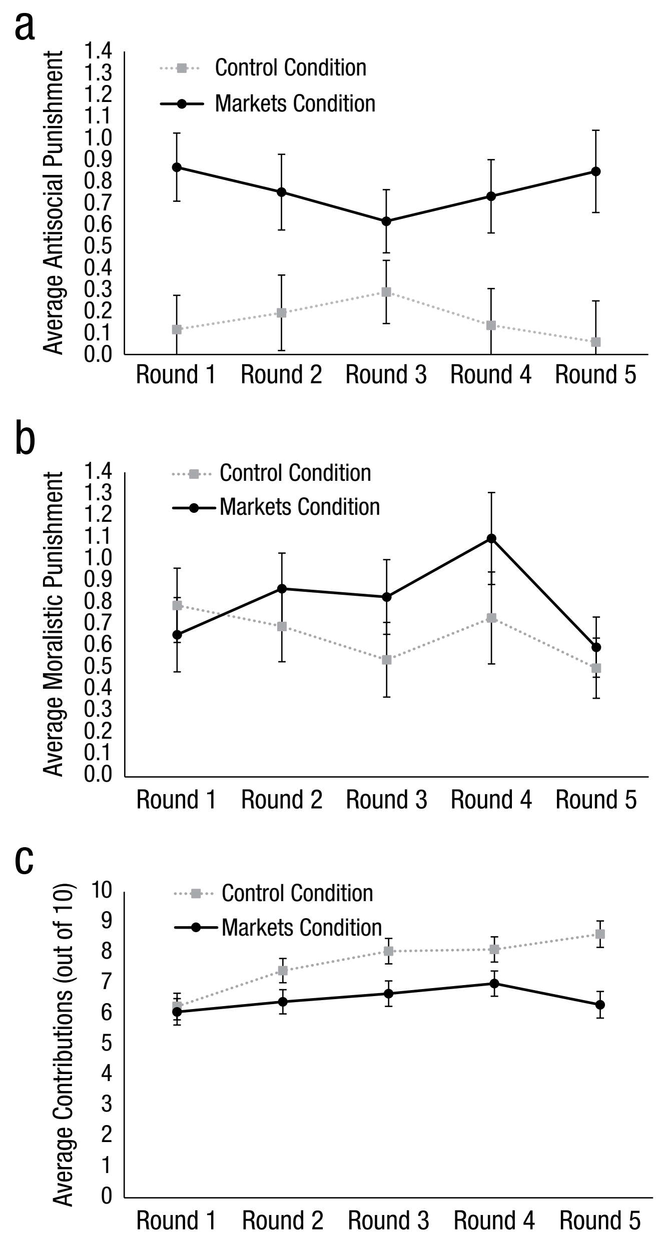

Fig. 1. Average per-person (a) antisocial punishment, (b) moralistic punishment, and (c) contributions to the public good in each round, separately for the control and markets conditions. Error bars represent standard errors of the mean.

comparison). To be certain, we had to eliminate three alternative explanations for the higher antisocial punishment in the markets condition: retaliation, differences in contributions, and confusion. First, antisocial punishment could have been high in the markets condition as a reaction to receiving moralistic punishment. However, this alternative cannot explain our main result, because there was more antisocial punishment in the markets condition than in the control condition even in the first round before participants received punishment from others. Furthermore, there was slightly less—not more—moralistic punishment in the first round of the markets condition than the control condition (M = L\$2.62, SE = L\$0.72 vs. M = L\$3.15, SE = 0.55, respectively), $t(24) = 0.59$, $p = .56$, $d = -0.23$, 95% CI for d = [−0.99, 0.55], which is the opposite of what we would have predicted if the high antisocial punishment in the markets condition was merely retaliation for moralistic punishment. Overall, the markets and control conditions did not differ in the amount of moralistic punishment per round (M = L\$3.23, SE = L\$0.44 vs. M = L\$2.60, SE = L\$0.44, respectively), $F(1, 24) = 1.03$, $p = .32$, $\eta_p^2 = .04$ (Fig. 1b); moralistic punishment did not differ across rounds, $F(4, 96) = 1.33$, $p = .26$, $\eta_p^2 = .05$, nor was there a Round × Condition interaction, $F(4, 96) = 0.70$, $p = .60$, $\eta_p^2 = .03$. This suggests that the higher antisocial punishment in the markets condition is not simply a reaction to receiving moralistic punishment. If anything, it might be the opposite: Some moralistic punishment might be retaliation for the first-round antisocial punishment. The two conditions differed in the type of punishment used—interaction $F(1, 24) = 6.89$, $p = .015$, $\eta_p^2 = .22$: There was much more moralistic punishment than antisocial punishment in the control condition, $F(1, 12) = 28.34$, $p < .001$, $\eta_p^2 = .70$, but the two types of punishment were not different in the markets condition, $F(1, 12) = 0.11$, $p = .75$, $\eta_p^2 = .01$.

A second possible alternative explanation is that the two conditions differed in contributions. However, the markets and control conditions did not differ in first-round contributions (M = L\$6.08, SE = L\$0.43 vs. M = L\$6.25, SE = L\$0.43, respectively), $t(24) = 0.28$, $p = .78$, $d = 0.11$, 95% CI for d = [−0.66, 0.88] (Fig. 1c); even the upper range of this confidence interval is much smaller than the effect size for antisocial punishment, $d = 1.41$. This suggests that the higher antisocial punishment in the first round of the markets condition was not the result of differential contributions. Interestingly, contributions did show a Round × Condition interaction, $F(2.75, 63.13) = 4.83$, $p = .005$, $\eta_p^2 = .17$: There was no effect of round on contributions in the markets condition, $F(4, 44) = 1.00$, $p = .42$, $\eta_p^2 = .08$, but there was an effect of round in the control condition, $F(4, 48) = 19.75$, $p < .001$, $\eta_p^2 = .62$; specifically, contributions increased linearly in the control condition, $F(1, 12) = 74.85$, $p < .001$, $\eta_p^2 = .86$. This suggests that the higher antisocial punishment in the markets condition was effective at preventing cooperating from escalating in the face of partner choice and that competition over partners can reduce cooperation if people compete by suppressing others' cooperation. However, given that any differences appeared only in later rounds, they cannot explain away the higher antisocial punishment that was present from Round 1.

A third possible alternative explanation is that participants were more confused in the markets condition because of the trust game that followed the PGG, and this confusion led to higher antisocial punishment. Indeed, some sessions contained participants who punished themselves, either by mistake or by misunderstanding, and these sessions tended to be in the markets condition rather than the control condition (6/13 vs. 1/13, respectively; $z = 2.48$, $p = .01$). However, even if we excluded any session in which participants punished themselves, there was still more antisocial punishment per round in the markets condition than in the control condition (M = L\$2.06, SE = L\$0.35 vs. M = L\$0.63, SE = L\$0.27), $F(1, 17) = 10.49$, $p = .005$, $\eta_p^2 = .38$, $d = 1.53$, 95% CI for d = [0.41, 2.51]. This conservative analysis shows that although there may have been more confusion in the markets condition, the higher antisocial punishment in that condition was not simply caused by a few confused participants.

Altogether, antisocial punishment was higher in the markets condition than in the control right from Round 1 and did not change over rounds. This difference cannot be explained by different contributions, by different amounts of moralistic punishment, or by confusion. Thus, the main explanation remaining is that the difference in antisocial punishment was caused by competition to be chosen to play the trust game.

Trust game (markets condition only)

In the markets condition, the PGG was followed by a trust game. Observers chose the highest contributor from the PGG in 10 of 11 groups where there was a single highest contributor (binomial $p < .001$ against chance probability of 1/4); two other groups had 2 participants tied for highest contributor, and in both of these groups, one of the top contributors was chosen. This indicates that contributions were perceived as signals of trustworthiness. Observers also trusted more money to players with higher average contributions, $r(11) = .28$, $p = .35$, 95% CI = [−.31, .72]; this was not statistically significant because of the low sample size, but it is comparable with findings of previous studies when people are competing to be chosen and is lower than when people are not competing to be chosen (Barclay & Willer, 2007). Participants with higher average contributions were willing to return a higher percentage in a hypothetical trust game, $r(50) = .27$, $p = .052$, 95% CI = [−.01, .51], and when told the amount entrusted for eliciting actual returns with the strategy method, they returned a higher actual amount, $r(44) = .42$, $p = .004$; controlling for amount entrusted: partial $r(43) = .42$, $p = .004$, 95% CI = [.17, .65]. These results replicate those of previous work showing that people trust cooperators and that such trust is broadly warranted.

Responses to questionnaires

We conducted additional analyses at the individual level, in which we treated each participant as his or her own data point.

Feelings of competition. Our manipulation check was successful: On a scale from 0 to 3, participants in the markets condition reported feeling that they were more in competition with their group members (M = 1.52, SE = 0.12) than did participants in the control condition (M = 1.21, SE = 0.08), unequal-variance $t(87.88) = 2.12$, p = .04, d = 0.42, 95% CI for d = [0.03, 0.81]. We also measured personality differences in cooperativeness and competitiveness using the CCPS: Participants in the markets condition were not significantly different from those in the control condition in either competitiveness (average score = 10.2, SE = 0.3 vs. average score = 10.0, SE = 0.3, respectively), $t(102)$ = 0.36, p=.72, d = 0.07, or in cooperativeness (average score = 11.6, SE = 0.3 vs. average score = 12.2, SE = 0.2, respectively), $t(102)$ = 0.36, p=.13, d = 0.30. Both the manipulation check and the CCPS were administered after the PGG and trust game, but only the temporary measure of competition differed between conditions (manipulation check), whereas the personality measure of competition did not (CCPS).

Correlations between personality and game behavior. Table 1 presents correlations among game behavior and the CCPS subscales. Antisocial punishment was negatively correlated with contributions in both the markets and control conditions, $r(50)$ = −.45 (95% CI = [−.64, −.20]) and $r(50)$ = −.51 (95% CI = [−.69, −.28]), respectively. Antisocial punishment was positively correlated with moralistic punishment in the markets condition, $r(50)$ = .39, p = .005, 95% CI = [.13, .60], but not in the control condition, $r(50)$ = .02, p = .90, 95% CI = [−.25, .29]; these correlations differed with z = 1.92 (two-tailed p = .055 because we did not predict this). Many participants carried out both moralistic and antisocial punishment in the same round: 24 people in the markets condition versus 8 in the control condition (see Table S1 in the Supplemental Material available online). Average contributions were significantly correlated with the cooperative subscale of the CCPS in the control condition, $r(50)$ = .47, p < .001, 95% CI = [.23, .66], but not in the markets condition, $r(50)$ = .12, p = .39, 95% CI = [−.16, .38]; these correlations differed with z = 1.91 (one-tailed p = .027), and the one-tailed test was justified because our hypothesis explicitly predicted that antisocial punishment would suppress cooperation in the markets condition (thus reducing the correlation between cooperative personality and contributions).

Exploratory analysis of antisocial punishment by loners

Rand and Nowak (2011) proposed that individuals who prefer a loner strategy in PGGs (i.e., prefer a fixed payoff for sitting out the PGG instead of playing it) are most likely to engage in antisocial punishment when forced to play the game. We asked our participants the hypothetical question of whether they would prefer a fixed payoff of L$12 (henceforth "loners") or another PGG with new partners (henceforth "nonloners"). In our sample, the 64 loners and 44 nonloners did not differ in their use of antisocial punishment (M = L$2.08,

Table 1. Bivariate Correlations (n = 52) Between Personality and Game Behavior

Variable	Average contribution	Moralistic punishment	Antisocial punishment	Cooperative score	Competitive score
Average contribution	—	−.053	−.51**	.47**[a]	.01
Moralistic punishment	− .17	—	.02[b]	−.03	.27†[c]
Antisocial punishment	−.45**	.39*[b]	—	−.15	.08
Cooperative score	.12[a]	−.04	−.02	—	.18
Competitive score	−.01	.21[c]	.22	.16	—
Percentage returned in hypothetical trust game	.27†	−.08	−.19	.06	.05
Amount returned in trust game[d]	.42*	−.22	−.20	.07	−.02

Note: Results for the control condition are above the diagonal, and results for the markets condition are below the diagonal.
[a]Correlations significantly differed between conditions (z = 1.91, one-tailed p = .028, two-tailed p = .056). [b]Correlations significantly differed between conditions (z = 1.92, one-tailed p = .027, two-tailed p = .055). [c]The correlation was significant if both conditions were combined: N = 104, r = .24, p = .015, 95% confidence interval = [.05, .41]. [d]This is a partial correlation controlling for amount entrusted. Data from 6 participants were missing because of a computer error, so n = 46, df = 43.
†p < .10. *p < .005. **p < .001.

SE = L$0.78 vs. *M* = L$2.50, *SE* = L$0.52, respectively), $t(102) = 0.46$, $p = .65$, $d = 0.09$, 95% CI = [−0.31, 0.47]. To eliminate any effects of partner choice, we reran the analysis within the control condition only. Among this reduced sample, the 25 loners used significantly less antisocial punishment (*M* = L$0.26, *SE* = L$0.14) than the 27 nonloners (*M* = L$1.36, *SE* = L$0.45), $t(28.42) = 2.34$, $p = .03$, $d = 0.63$, 95% CI for the mean difference = [0.14, 2.06], 95% CI for *d* = [0.05, 1.17]; this is in the opposite direction predicted by Rand and Nowak. Thus, we failed to replicate Rand and Nowak's findings, and interestingly, found the opposite pattern of expected behavior when analyzing only the control group's data.

Discussion

Barclay (2013) and Sylwester and colleagues (2013) proposed that when there is competition to be chosen as a social partner, low cooperators will use antisocial punishment or do-gooder derogation to reduce others' cooperation and make themselves look good by comparison. Our main result supports this: Antisocial punishment was much higher in the markets condition than in the control condition, with a very strong effect size ($d = 1.51$). This difference was present in the first round. This effect cannot be explained by differences in contributions because contributions were not different in the first round (consistent with previous research; Barclay, 2004). Although some instances of antisocial punishment may be retaliation for moralistic punishment (Herrmann et al., 2008; Sylwester et al., 2013; Van Dijk et al., 2015), our main effect cannot be explained as simply retaliation because moralistic punishment did not differ between the two conditions either in the first round or overall. Furthermore, these results persisted even when we eliminated sessions in which participants may have been confused. Thus, it appears to be competition over partners that causes the higher antisocial punishment.

Previous studies have shown an escalation of cooperation when people compete to be chosen as partners (Barclay & Willer, 2007; Sylwester & Roberts, 2010). This was not the case in this experiment. In the control condition, contributions increased across rounds, as usually happens when moralistic punishment is present (see the meta-analysis by Balliet et al., 2011). However, contributions did not increase in the markets condition. This is exactly what we would predict if antisocial punishment was a strategy to bring down the high cooperators and prevent them from competing to cooperate the most: Antisocial punishment prevents an escalation of generosity (Barclay, 2013). This result is also consistent with previous work showing that high levels of antisocial punishment tend to inhibit cooperation and undermine the effects of moralistic punishment (Fatas & Mateu, 2015; Herrmann et al., 2008; Rand & Nowak, 2011). Furthermore, cooperative personalities (on the CCPS) were less related to contributions in the markets condition, which is also consistent with antisocial punishment having suppressed people's generosity. These results suggest that contrary to previous suggestions, biological markets do not always produce good outcomes and can even produce bad outcomes, because organisms may also compete to be chosen using harmful tactics.

When will it pay to compete via antisocial punishment and do-gooder derogation? All else being equal, these strategies will pay better when there is more competition, especially if the derogator and target are competing most directly, such that one's gain is a relatively bigger loss for the other. Antisocial punishment and do-gooder derogation are risky tactics because they could make oneself look bad if done poorly. Thus, any factors that affect the payoff for risk taking should affect the payoffs for antisocial punishment. These include (Barclay, Mishra, & Sparks, 2018; Mishra, Barclay, & Sparks, 2017) the probability that punishment or derogation will affect the target's desirability (i.e., probability of success), the amount by which it would do so (i.e., benefits if successful), the amount it will hurt the derogator to be seen to punish or derogate (i.e., the cost if unsuccessful), and the derogator's likelihood of being chosen as a partner without resorting to this risky strategy (i.e., the level of desperation). Future studies could investigate the effects of each of these on antisocial punishment and do-gooder derogation.

One limitation is that we used laboratory games to assess cooperation and antisocial punishment instead of real-world behavior. However, such games are commonly used as models of real-world phenomena, with the monetary incentives representing the real-world costs and benefits of helping or harming others. Antisocial punishment in PGGs is a good model for understanding do-gooder derogation, which is where people who help others or act morally receive criticism or ridicule (Minson & Monin, 2012; Monin, 2007) or are evicted from their groups (Parks & Stone, 2010). Another limitation is that punishers' identities were not made visible. Unjustified punishment is viewed as untrustworthy (Barclay, 2006), so antisocial punishment might have been more muted or subtle if the observer had known who had punished whom. Outside the laboratory, antisocial punishment and do-gooder derogation are likely done subtly or indirectly to reduce the chance of looking bad. Our study demonstrates people's motivations for engaging in antisocial punishment when such constraints are absent and shows that introducing partner competition makes people more willing to engage in antisocial punishment.

Overall, our results support the idea that antisocial punishment and do-gooder derogation function as a way to prevent one's competitors from gaining relative reputation, which would make oneself look worse by comparison (Barclay, 2013). These results can apply to any social dilemma or cooperative situation, including organizations, corporations, group projects, hunter-gatherer food sharing, resource conservation, and philanthropy. For example, our results may explain why very generous people sometimes hide large philanthropic donations (Raihani, 2014): Others may resent such actions and punish or ostracize the donors. Our study suggests that perhaps the best strategy for social competition in any of these areas is to be good but not *too* good, lest one make enemies of those who are *merely* good and thus receive their punishment.

Action Editor

Eddie Harmon-Jones served as action editor for this article.

Author Contributions

A. Pleasant programmed, conducted, and codesigned the study; performed the primary statistical analyses; and wrote the first draft of the manuscript. P. Barclay conceived and designed the study; sought funding; supervised the programming, running, and analysis of the study; conducted some secondary statistical analyses; wrote the submitted draft of the manuscript; and navigated the submission process, including all revisions.

Acknowledgments

We thank the Barclay Lab for Experimental Evolutionary Psychology (Hayley Asling, Savreet Chuckal, Sara Kafashan, Megan Kelly, Andrea Larney, Jillian O'Connor, Amanda Rotella, Clara Sankey) for discussions.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This research was funded by the Social Sciences and Humanities Research Council of Canada.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617752642

Open Practices

All data and materials have been made publicly available via Dryad and can be accessed at https://datadryad.org//resource/doi:10.5061/dryad.q88sm. The design and analysis plans were not preregistered. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617752642. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

References

Balliet, D., Mulder, L., & Van Lange, P. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin*, *137*, 594–615. doi:10.1037/a0023489

Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the "tragedy of the commons." *Evolution & Human Behavior*, *25*, 209–220.

Barclay, P. (2006). Reputational benefits for altruistic punishment. *Evolution & Human Behavior*, *27*, 325–344.

Barclay, P. (2011). Competitive helping increases with the size of biological markets and invades defection. *Journal of Theoretical Biology*, *281*, 47–55.

Barclay, P. (2012). Proximate and ultimate causes of strong reciprocity and punishment. *Behavioral & Brain Sciences*, *35*, 16–17.

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution & Human Behavior*, *34*, 164–175. doi:10.1016/j.evolhumbehav.2013.02.002

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology*, *7*, 33–38. doi:10.1016/j.copsyc.2015.07.012

Barclay, P., Mishra, S., & Sparks, A. M. (2018). *Relative state creates individual differences in risk-taking*. Manuscript submitted for publication.

Barclay, P., & Van Vugt, M. (2015). The evolutionary psychology of human prosociality: Adaptations, mistakes, and byproducts. In D. Schroeder & W. Graziano (Eds.), *Oxford handbook of prosocial behavior* (pp. 37–60). Oxford, England: Oxford University Press.

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, *274*, 749–753. doi:10.1098/rspb.2006.0209

Barr, A. (2001). *Social dilemmas and shame-based sanctions: Experimental results from rural Zimbabwe* (Working Paper WPS/2001-11). Oxford, England: Centre for the Study of African Economies.

Boehm, C. (1999). *Hierarchy in the forest: The evolution of egalitarian behavior*. Cambridge, MA: Harvard University Press.

Buss, D. M., & Dedden, L. A. (1990). Derogation of competitors. *Journal of Social and Personal Relationships*, *7*, 395–422.

Cinyabuguma, M., Page, T., & Putterman, L. (2006). Can second-order punishment deter perverse punishment? *Experimental Economics*, *9*, 265–279. doi:10.1007/s10683-006-9127-z

Delton, A., Nemirow, J., Robertson, T., Cimino, A., & Cosmides, L. (2013). Merely opting out of a public good is moralized: An error management approach to cooperation. *Journal of Personality and Social Psychology*, *105*, 621–638. doi:10.1037/a0033495

Fatas, E., & Mateu, G. (2015). Antisocial punishment in two social dilemmas. *Frontiers in Behavioral Neuroscience*, *9*, Article 107. doi:10.3389/fnbeh.2015.00107

Fehr, E., Hoff, K., & Kshetramade, M. (2008). Spite and development. *The American Economic Review*, *98*, 494–499.

Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science*, *25*, 656–664. doi:10.1177/0956797613510184

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*, 171–178. doi:10.1007/s10683-006-9159-4

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sundue, U., & Elger, C. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, *318*, 1305–1308.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, *53*, 3–35. doi:10.1016/S0167-2681(03)00094-5

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*, 1362–1367.

Kuběna, A. A., Houdek, P., Lindová, J., Příplatová, L., & Flegr, J. (2014). Justine effect: Punishment of the unduly self-sacrificing cooperative individuals. *PLOS ONE*, *9*(3), Article e92336. doi:10.1371/journal.pone.0092336

Lu, S., Au, W., Jiang, F., Xie, X., & Yam, P. (2013). Cooperativeness and competitiveness as two distinct constructs: Validating the Cooperative and Competitive Personality Scale in a social dilemma context. *International Journal of Psychology*, *48*, 1135–1147. doi:10.1080/00207594.2012.743666

Minson, J. A., & Monin, B. (2012). Do-gooder derogation: Disparaging morally motivated minorities to diffuse anticipated reproach. *Social Psychological & Personality Science*, *3*, 200–207. doi:10.1177/1948550611415695

Mishra, S., Barclay, P., & Sparks, A. (2017). The relative state model: Integrating need-based and ability-based pathways to risk-taking. *Personality and Social Psychology Review*, *21*, 176–198. doi:10.1177/1088868316644094

Monin, B. (2007). Holier than me? Threatening comparison in the moral domain. *Revue Internationale de Psychologie Sociale*, *20*, 53–68.

Nikiforakis, N., & Normann, H. (2008). A comparative statics analysis of punishment in public good experiments. *Experimental Economics*, *11*, 358–369. doi:10.1007/s10683-007-9171-3

Noë, R., & Hammerstein, P. (1994). Biological markets: Supply and demand determine the effect of partner choice in cooperation, mutualism and mating. *Behavioral Ecology and Sociobiology*, *35*, 1–11. doi:10.1007/BF00167053

Noë, R., & Hammerstein, P. (1995). Biological markets. *Trends in Ecology & Evolution*, *10*, 336–339.

Parks, C., & Stone, A. (2010). The desire to expel selfish members from the group. *Journal of Personality and Social Psychology*, *99*, 303–310. doi:10.1037/a0018403

Prediger, S., Vollan, B., & Herrmann, B. (2014). Resource scarcity and antisocial behavior. *Journal of Public Economics*, *119*, 1–9. doi:10.1016/j.jpubeco.2014.07.007

Raihani, N. J. (2014). Hidden altruism in a real-world setting. *Biology Letters*, *10*(1), Article 20130884. doi:10.1098/rsbl.2013.0884

Raihani, N. J., & Smith, S. (2015). Competitive helping in online giving. *Current Biology*, *25*, 1183–1186. doi:10.1016/j.cub.2015.02.042

Rand, D. G., & Nowak, M. A. (2011). The evolution of antisocial punishment in optional public goods games. *Nature Communications*, *2*, Article 434. doi:10.1038/ncomms1442

Roberts, G. (1998). Competitive altruism: From reciprocity to the handicap principle. *Proceedings of the Royal Society B: Biological Sciences*, *265*, 427–431.

Scott-Phillips, T. C., Dickins, T. E., & West, S. A. (2011). Evolutionary theory and the ultimate-proximate distinction in the human behavioral sciences. *Perspectives on Psychological Science*, *6*, 38–47.

Sylwester, K., Herrmann, B., & Bryson, J. J. (2013). Homo homini lupus? Explaining antisocial punishment. *Journal of Neuroscience, Psychology, and Economics*, *6*, 167–188. doi:10.1037/npe0000009

Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. *Biology Letters*, *6*, 659–662. doi:10.1098/rsbl.2010.0209

Tinbergen, N. (1968). On war and peace in animals and man. *Science*, *160*, 1411–1418.

Van Dijk, E., Molenmaker, W., & Kwaadsteniet, E. (2015). Promoting cooperation in social dilemmas: The use of sanctions. *Current Opinion in Psychology*, *6*, 118–122. doi:10.1016/j.copsyc.2015.07.006

Voltaire. (1865). *La bégueule. In Romans de Voltaire: Suivis de ses contes en vers* [Novels of Voltaire: Followed by his tales in verse]. Paris, France: Garnier Frères. (Original work published 1772)

Xie, X., Yu, Y., Chen, X., & Chen, X. P. (2006). The measurement of cooperative and competitive personality. *Acta Psychologica Sinica*, *38*, 116–125.

Zahavi, A., & Zahavi, A. (1997). *The handicap principle: A missing piece of Darwin's puzzle*. New York, NY: Oxford University Press.