European Journal of Personality, *Eur. J. Pers.* **29**: 498–505 (2015)
Published online 19 May 2015 in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/per.2003

Understanding the Dark Side of Costly Punishment: The Impact of Individual Differences in Everyday Sadism and Existential Threat

STEFAN PFATTHEICHER[1]* and SIMON SCHINDLER[2]

[1]*Department of Social Psychology, Ulm University, Germany*
[2]*Department of Social Psychology, Kassel University, Germany*

Abstract: In public goods situations, a specific destructive behaviour reliably emerges when individuals face the possibility of costly punishing others: antisocial punishment, that is, costly punishing cooperative individuals. So far, however, little is known about the individual differences and situational factors that are associated with the dark side of costly punishment. This research deals with this shortcoming. We argue that antisocial punishment reflects the basic characteristics of sadism, namely, aggressive behaviour to dominate and to harm other individuals. We further argue that antisocial punishment may reflect a type of behaviour that allows for the maintenance of self-esteem (through aggressively dominating others). Therefore, we expect that individuals who report a disposition for everyday sadism are particularly likely to engage in antisocial punishment when their self has been threatened (by thinking about one's own death). In a study (N = 99), we found empirical support for this assumption. The present research contributes to a better understanding of antisocial punishment and suggests that sadistic tendencies play a crucial role, especially when the self is (existentially) threatened. Copyright © 2015 European Association of Personality Psychology

Key words: antisocial punishment; dominance; mortality salience; sadism; social dilemma

INTRODUCTION

Imagine a (too) highly motivated person in a student learning group who is, however, socially disapproved by the other group members. Think about so-called 'do-gooders', for instance vegetarians, who are sometimes faced with hostility and derogation (Minson & Monin, 2012, Monin, Sawyer, & Marquez, 2008). These are examples in which individuals doing good things are harmed by others. The present research examines one similar behaviour that emerges in social dilemma situations: antisocial punishment, that is, using one's own resources to costly punish cooperative individuals.

In social dilemma situations, individuals are better off when they behave uncooperatively rather than cooperatively (cf. van Lange, Balliet, Parks, & van Vugt, 2014, van Lange, Joireman, Parks, & van Dijk, 2013). For instance, during a drought, farmers can overuse water to save their own harvest (cf. Ostrom, 1990). The problem with this is that saving one's own harvest by overusing water damages the harvests of other farmers. Thus, the problem (i.e. the dilemma) is that an individual benefits from uncooperative behaviour that results, however, in a reduced benefit for the collective. This problem is inherent in social dilemma situations and ultimately results in the question of how uncooperative behaviour can be avoided. One prominent solution in this regard is to establish a system of costly punishment, that is, the option to invest private resources to punish interaction partners (Fehr & Gächter, 2002). In fact, a remarkable amount of empirical evidence (cf. the meta-analysis of Balliet, Mulder, & van Lange, 2011) has convincingly shown that (i) punishing uncooperative others leads to an increase in cooperative behaviour of the punished individuals in future interactions and (ii) there is a higher cooperation level in situations in which an option to punish is available as compared with situations in which no such option lable. Interestingly, in such situations, it is not only uncooperative individuals that are punished but also cooperative individuals (i.e., antisocial punishment; Herrmann, Thöni, & Gächter, 2008).

Although antisocial punishment reliably emerges across societies (Herrmann, Thöni, & Gächter, 2008), it can be seen when reviewing the literature on antisocial punishment that only a few empirical papers have investigated this interesting behavioural phenomenon (cf. Sylwester, Herrmann, & Bryson, 2013). That is, the investigation of antisocial punishment has been almost completely neglected in social dilemma research thus far (cf. Dreber & Rand, 2012, Herrmann et al., 2008). Published empirical papers mainly focus on the boundary conditions of antisocial punishment. For instance, Herrmann et al. (2008) showed that antisocial punishment is more likely to emerge in societies with relatively weak norms of civic cooperation and a non-established and malfunctioning judiciary (see also Gächter & Herrmann, 2009). Bernhard, Fischbacher and Fehr (2006) document (on a descriptive level) that antisocial punishment is more likely to emerge when an out-group member (as compared with an in-group member) can be punished (for similar findings, see Goette, Huffman, Meier, & Sutter, 2012).

An investigation that examines psychological factors of antisocial punishment is almost completely lacking. An exception is the research by Parks and Stone (2010) and Pfattheicher, Landhäußer and Keller (2014). Parks and Stone (2010) showed that cooperative group members establish an undesirable high

*Correspondence to: Stefan Pfattheicher, Universität Ulm, Abteilung Sozialpsychologie, 89069, Ulm, Germany.
E-mail: stefan.pfattheicher@uni-ulm.de

Received 23 September 2014
Revised 30 January 2015, Accepted 11 April 2015

standard for all group members, leading to the tendency to ostracize cooperative group members (see also Irwin & Horne, 2013). Social ostracism is some sort of punishment (Feinberg, Willer, & Schultz, 2014). A motivational perspective was applied by Pfattheicher et al. (2014) who argued that antisocial punishment reflects aggressive behaviour to dominate and to harm other individuals (see also Sylwester et al., 2013, who made the same argument). Accordingly, Pfattheicher et al. (2014) showed that antisocial punishment could be predicted by aggressive dominance concerns and the dominance-related hormone testosterone given a relatively low level of cortisol (i.e., the dual-hormone hypothesis; Carré & Mehta, 2011, Mehta & Josephs, 2010).

Still, the analysis of antisocial punishment is in its infancy. In the present work, therefore, we investigate what personal and situational factors are associated with an individual's tendency to engage in antisocial punishment. Specifically, we focus on the interplay of individual differences in everyday sadism and existential threats to the self (i.e., mortality salience, MS). In this sense, our approach to examining the associations between these factors and antisocial punishment reflects a step in a new direction and constitutes a contribution to the understanding of antisocial punishment in social dilemma situations that addresses a gap in this field of study. The following outlines the rationale for why we expect sadism and existential threats to the self to be crucial in engaging in antisocial punishment.

ANTISOCIAL PUNISHMENT AND SADISM

In the present work, we argue that everyday sadists are particularly prone to engaging in antisocial punishment under specific conditions. The very essence of sadism is that sadists are motivated to dominate and to control other individuals by harming them because they experience pleasure through their cruelty (Cooke, 2001, Dietz, Hazelwood, & Warren, 1990, O'Meara, Davies, & Hammond, 2011). So far, sadism has principally been investigated in a clinical context and in descriptive analyses portraying individuals who experience pleasure when hurting and dominating others (e.g., Baumeister & Campbell, 1999, Baumeister & Vohs, 2004, Fromm, 1973, Piven, 2003). Nonetheless, sadistic actions also emerge in the subclinical population (Chabrol, Van Leeuwen, Rodgers, & Séjourné, 2009). Buckels, Jones and Paulhus (2013) have documented that everyday people vary in their sadistic tendencies and showed that those with a relatively strong (as compared with a weak) sadistic disposition were more likely to engage in killing bugs and harming an innocent person, particularly when personal costs had to be incurred. Greitemeyer (2015) reports that everyday sadism predicts the amount of violent video gameplay. Moreover, Buckels, Trapnell and Paulhus (2014) document that everyday sadists are more likely to engage in trolling, a practice that reflects evil and destructive behaviour towards innocent others on the Internet. In the same vein, Reidy, Zeichner and Seibert (2011) showed that the faster an individual responded to cruel images after a positive word was presented (i.e., implicit sadism), the more likely they were to engage in unprovoked aggression. This research in fact shows that everyday sadists exist and that these individuals engage in costly, harmful actions that control other individuals' personal states.

Taking up the notion of antisocial punishment, we argue that the punishment of cooperative others, which reflects aggressive behaviour to dominate and to harm other individuals (Pfattheicher et al., 2014, Sylwester et al., 2013), fits the evil mould of sadism. Thus, we assumed that everyday sadists are prone to engaging in antisocial punishment. We expect that everyday sadists engage in antisocial punishment particularly when their self has been threatened. The rationale for this assumption is outlined in the following section.

EXISTENTIAL THREAT AND HARMING OTHERS

Humans have a fundamental psychological need to believe that they are a valuable being and therefore seek to have positive self-esteem (Pyszczynski, Greenberg, Solomon, Arndt, & Schimel, 2004). When the self is threatened, individuals are motivated to restore their self-esteem (e.g., Leary, 1999, Leary & Baumeister, 2000), for instance, when the self is existentially threatened by the recognition of one's own mortality (Greenberg, Solomon, & Arndt, 2008). Indeed, one main assumption of terror management theory (TMT) is that humans have a fundamental psychological need for self-esteem to cope with the potentially paralyzing threat of one's own mortality (Pyszczynski et al., 2004).

One way to maintain self-esteem after the self has been threatened is to engage in harming and dominating others (Baumeister, Bushman, & Campbell, 2000, Baumeister, Smart, & Boden, 1996, Fein & Spencer, 1997, Leary, 1999). In line with this perspective, Brown and Zeigler-Hill (2004) documented a positive relation between individual differences in dominance and self-esteem. Applying a causal analysis approach, Leary, Cottrell and Phillips (2001) showed that strengthening one's perceived dominance in a leadership context increased self-esteem. Moreover, when the self was threatened via a social exclusion paradigm, as compared with a no-threat condition, individuals were more likely to engage in aggression towards an offender as well as an innocent third person (Twenge, Baumeister, Tice, & Stucke, 2001, Twenge & Campbell, 2003). In this regard, Baumeister et al. (1996) wrote that 'a successful violent attack achieves a symbolic dominance over the other person, and so it affirms one's esteem to the extent of being superior to the victim' (p. 11). Building on these considerations, we argue that harming and dominating cooperative others reflect one strategy to feel superior and to maintain one's self-esteem (cf. Crocker, Lee, & Park, 2004, Crocker & Park, 2004).

In sum, in the present work, we build on the following: (i) the theoretical account of TMT postulating that individuals are motivated to maintain their self-esteem after being existentially threatened (Pyszczynski et al., 2004) and (ii) research showing that sadistic behaviours, like harming and aggressively dominating others, serve a self-esteem enhancing function (Baumeister et al., 1996, Crocker & Park, 2004, Crocker et al., 2004). Specifically, assuming antisocial punishment to be related to sadistic motives, we expected individuals

DOI: 10.1002/per

with a disposition for sadistic tendencies to show increased engagement in antisocial punishment especially when being confronted with an existential threat (i.e., MS).

In contrast, we do not expect that individual differences in sadism predict the punishment of uncooperative free riders (altruistic punishment) because the motives underlying altruistic punishment reflect equality concerns (Fowler, Johnson, & Smirnov, 2004), group orientation (Barclay, 2006) and fairness concerns (Fehr & Fischbacher, 2003, Singer & Steinbeis, 2009, Tabibnia, Satpute, & Lieberman, 2008). Conceptually, these motives are not closely related to sadism. One may also assume that individual differences in sadism do not predict altruistic punishment because there is nothing specific to sadism in altruistic punishment as non-sadists (because of the motives mentioned earlier) are also likely to engage in altruistic punishment.

Although not in the focus of the present contribution, we do, however, expect that existential threat increases people's motivation to punish uncooperative others (altruistic punishment). This should hold true because individuals defend important values and social norms (e.g., Rosenblatt, Greenberg, Solomon, Pyszczynski, & Lyon, 1989). The defence of the social norms of fairness and equality might be possible by altruistic punishment (Fehr & Fischbacher, 2003, Fowler et al., 2004, Singer & Steinbeis, 2009, Tabibnia et al., 2008). To test this assumption, we also include the punishment of uncooperative free riders in our analyses, as reported in the succeeding text.

METHOD

In this study, we measured individuals' dispositions to everyday sadism (Buckels et al., 2013, Paulhus & Jones, in press) and existentially threatened the individuals' selves via an MS manipulation (Burke, Martens, & Faucher, 2010, Greenberg et al., 1990) after participants had read the explanation of a public goods game. After the existential threat or a control treatment occurred, individuals played a typical public goods game with the option to costly punish other group members (cf. Fehr & Gächter, 2002).

Sample

The study consisted of 99 University of Ulm students ($M_{age} = 21.09$; 46.41% women).

Sadism

Dispositional sadism was measured using six items of the Varieties of Sadistic Tendencies Scale, which directly assess sadism (Buckels et al., 2013, Paulhus & Jones, in press).[1] The scale endpoints of the items were labelled (1) *not at all true* and (7) *completely true* ($\alpha = .67$, $M = 2.04$, $SD = 0.82$).

[1]The full scale consists of seven items. We did not assess the item 'I enjoy hurting my partner during sex (or pretending to)' because of ethical reasons. Moreover, we ran seven sessions of the public goods game. In three sessions, dispositional sadism was assessed before the explanation of the public goods game was read by participants; in four sessions, it was assessed after participants played the public goods game. Order did not moderate the effects ($p = .79$).

Existential threat

Participants were randomly assigned to an MS condition or a dentist-visit control condition. Participants in the MS condition ($n = 47$) answered two open-ended questions about death, while those in the dentist-visit condition ($n = 52$) were asked two questions about a visit to the dentist (cf. Jonas et al., 2008). When using such explicit death primes, a distractor is necessary to diminish consciousness of death (Arndt, Greenberg, & Cook, 2002). Therefore, as in many studies on TMT (Burke et al., 2010), participants filled out the 20-item Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988, Krohne, Egloff, Kohlmann, & Tausch, 1996) and a five-item distractor questionnaire about sleep and waking patterns (Fritsche, Jonas, & Fankhänel, 2008).

Public goods game

We investigated antisocial punishment using a standard public goods game with the costly option to punish (cf. Fehr & Gächter, 2002, Pfattheicher & Keller, 2014).[2] As in a typical public goods game, four players constituted one group (cf. Fehr & Gächter, 2002). Each player was endowed with 20 money units (MUs; one MU was equal to €1 cent ~ $1.08 cent) and was free to choose how many of these to keep and how many to contribute to a common group project (i.e., the public good). Each MU contributed was multiplied by 1.6. Next, each player received one-fourth of the public good, independent of his or her contribution. Accordingly, it was always in the material self-interest of every individual to keep all of his or her MUs irrespective of how much the other three subjects contribute to the group project: if every group member invested 20 MUs, each subject would earn $(80 \times 1.6)/4$ MUs, that is, 32 MUs. If one group member engaged in free riding (e.g., he or she contributes 0 MU) and the other three group members still invest their 20 MUs, the free rider earns 44 MUs (20 MUs already owned plus one-fourth of the public good, that is, 24 MUs), and each of the other three group members earns 24 MUs.

Following this stage, each player was given information on the contributions made by the other three players and was then given the option to punish them by investing their own MUs (between 0 and 10 for each player), which reduced the selected other player's pay-off by a factor of three (e.g., the investment of two MUs decreases the pay-off of another by six MUs).[3]

[2]The instructions are provided in the supplementary material of Herrmann et al. (2008). We used exactly these instructions (as well as the same software).

[3]In this study, we predetermined the contributions of the other three players in each period to eliminate the variability that comes into play when the game involves real interactions. So we held the contributions of the other participants constant. The contributions of the other players were randomly drawn from a previous study involving real interactions (Pfattheicher & Keller, 2013), thus representing real behaviour. The contributions of the other three players were as follows: period 1: 5, 16, 20; period 2: 6, 9, 16; period 3: 2, 20, 14; period 4: 16, 0, 14; period 5: 15, 4, 15; and period 6: 20, 8, 20. How much participants were punished in each period was determined using previous data involving real interactions. If participants contributed less than six MUs, they were punished with 3-2-4-3-2-4 MUs (three MUs in the first period, two MUs in the second and so on). If participants contributed between 6 and 13 MUs, they were punished with 2-1-3-2-1-3 MUs, and if participants contributed more than 13 MUs, they were punished with 1-0-2-1-0-2 MUs (to realize antisocial punishment).

DOI: 10.1002/per

Six periods of the public goods game were played under anonymous conditions. All interactions were computer-mediated using z-Tree (Fischbacher, 2007), and all decisions were made simultaneously. Participants were told that the group composition changed from period to period to exclude direct reciprocity accounts (Trivers, 1971). Participants were paid their earnings privately ($M = €1.25 \sim \$1.74$, $SD = 0.17$) at the end of the session (and also received a chocolate bar or a beverage).

Antisocial punishment

In line with Herrmann et al. (2008), antisocial punishment was computed by summarizing the MUs across the six periods that were used by each player for the punishment of other players who contributed as much or more than the player himself or herself.

Altruistic punishment

The punishment of uncooperative individuals (altruistic punishment) was computed by summarizing the MUs across the six periods that were used by each player for the punishment of other players who contributed less than the player himself or herself.

RESULTS

Preliminary results

In order to give the reader an impression of the contributions to the public good and investments in punishment, the descriptive statistics are reported first. The mean number of MUs contributed to the public good across the six periods (maximum 6×20 MUs $= 120$ MUs) was 68.48 MUs ($SD = 29.49$). The mean number of MUs invested in antisocial punishment across the six periods was 3.02 MUs ($SD = 8.30$). Thirty-four out of 99 participants (34.34%) engaged in antisocial punishment. Contribution to the public good and investment in antisocial punishment were negatively correlated ($r = -.19$, $p = .06$). The mean number of MUs invested in the punishment of uncooperative individuals across the six periods was 6.45 MUs ($SD = 9.21$). Seventy-two out of 99 participants (72.72%) engaged in the punishment of uncooperative individuals at least once. As in the study of Fehr and Gächter (2002), contribution to the public good and investment in the punishment of uncooperative individuals were positively correlated ($r = .25$, $p = .01$). Antisocial punishment and the punishment of uncooperative individuals were also positively correlated ($r = .31$, $p < .01$) as in previous research (Pfattheicher et al., 2014). The Positive and Negative Affect Schedule values were not significantly affected by the manipulation ($ps > .91$), neither was the contribution to the public good ($p = .41$).

Main results

We report standard ordinary least squares regression and also Tobit regression to account for participants scoring zero on antisocial punishment (cf. McDonald & Moffitt, 1980, Pfattheicher & Keller, 2013). We also applied bootstrapping (based on 1000 resamples) to test for statistical robustness and to apply a non-parametric statistical test (Hayes, 2013). The results are displayed in Table 1. The findings revealed a significant sadism × condition interaction. Decomposing this interaction (Aiken & West, 1991, Preacher, Rucker, & Hayes, 2007) revealed a significant effect of the existential threat manipulation, but only when an individual's disposition for sadism was relatively strong (Table 1). That is, those with relatively strong sadistic tendencies engaged in antisocial punishment when their self was existentially threatened (as compared with the control condition), which supports our hypothesis. We want to note that this central finding remains robust in terms of significance levels and strength when controlling for the punisher's contribution, the amount of received punishment and altruistic punishment (see Supporting Information for statistical parameters).

In contrast, we did not find significant relations of individual differences in sadism and investment in altruistic punishment, in either the control or experimental conditions ($ps > .18$). Congruently, no significant interaction between sadism and condition emerged ($p = .92$; for more focused analyses and statistical parameters, see Supporting Information). We found, however, a main effect of the existential threat manipulation on the punishment of uncooperative individuals. Specifically, participants in the MS condition

Table 1. Regression coefficients of the main analyses ($N = 99$)

	OLS regression			Tobit regression
Criterion: antisocial punishment	*B* (*SE B*)	β	Bootstrapped 95% confidence interval	Coefficient (*SE*)
Constant	1.99 (1.05)		[0.77, 3.44]	−10.99 (3.41)***
Condition	1.91 (1.53)	0.12	[−0.72, 5.44]	6.38 (3.93)
Sadism	−0.73 (1.44)	−0.72	[−2.63, 1.12]	−1.48 (3.87)
Sadism × condition	6.33 (1.90)**	0.47	[1.60, 10.53]	10.78 (4.79)*
Effect of the condition at relatively low sadism	−3.25 (2.17)	−0.20	[−6.33, 0.05]	−2.41 (5.54)
Effect of the condition at relatively high sadism	7.08 (2.18)**	0.43	[1.21, 13.32]	15.17 (5.53)**

Note: OLS, ordinary least squares; SE, standard error. Sadism was mean centred; relatively low sadism refers to 1 standard deviation below mean; relatively high sadism refers to 1 standard deviation above mean; control condition = 0, mortality salience condition = 1.
*$p < .05$; **$p < .01$; ***$p < .01$.

DOI: 10.1002/per

punished more strongly ($M = 8.85$, $SD = 11.64$) compared with those in the control condition ($M = 4.29$, $SD = 5.53$; $t(64) = 2.45$, $p = .02$; corrected *df*).

DISCUSSION

In 1938, the French Nobel Prize-winning author Albert Camus started writing *Caligula*. In this play, Prince Caligula is confronted with the death of his beloved sister and mistress, which terrorizes his mind. After this disturbing event, Caligula reveals his evil side by murdering and bringing suffering upon those close to him. In the present paper, we have, in fact, delivered empirical data that support the plot of Camus' play. Specifically, we have shown that individuals who report a disposition for sadistic behaviours are particularly likely to engage in evil behaviour (i.e., antisocial punishment) after having been existentially threatened.

Pfattheicher et al. (2014) have argued that antisocial punishment reflects aggressive behaviour to dominate and to harm other individuals (Sylwester et al., 2013). In line with these findings, we were able to document that those individuals who are motivated to dominate and control others by harming them, that is, individuals who report a disposition for sadism, are the ones who engage in antisocial punishment after being existentially threatened. This finding in fact strengthens the conceptualization of antisocial punishment as a type of behaviour executed to dominate and to harm others. As such, the present work contributes to a better understanding of antisocial punishment, which is important given that research on social dilemma situations has almost completely neglected antisocial punishment thus far (Dreber & Rand, 2012, Herrmann et al., 2008).

From a social dilemma perspective, it is important to consider the conditions under which costly punishment of cooperative others may emerge. Proposing a system of costly punishment in social dilemma situations seems to be, at first sight, a worthwhile approach given its positive effects on the level of cooperation in social dilemma situations (Balliet et al., 2011). However, as shown in this work and in the work of Herrmann et al. (2008), antisocial punishment reliably emerges when a system of costly punishment is implemented. As such, knowing who engages in antisocial punishment and under what conditions can foster attempts to prevent and to deal with antisocial punishment.

The results obtained are also relevant from the perspective of TMT (Greenberg et al., 2008), in particular regarding the impact of MS on antisocial behaviour (cf. Schindler, Reinhard, & Stahlberg, 2013). We have documented that antisocial punishment in everyday sadists emerges after they are reminded of their own mortality, which supports the idea that antisocial punishment may be used by certain people (e.g., sadists) as a self-esteem-enhancing strategy to counter the fear of death. This is in line with the notion that harming and dominating others can reflect a strategy to feel superior and to maintain one's self-esteem (e.g., Baumeister et al., 1996, Baumeister et al., 2000, Brown & Zeigler-Hill, 2004, Crocker & Park, 2004, Crocker et al., 2004). Additionally, we could document that MS increases the punishment of uncooperative individuals (altruistic punishment). This finding is in line with TMT research (e.g., Rosenblatt et al., 1989) showing that—besides striving for self-esteem—existential threat increases people's motivation to defend important values and social norms (i.e., in our case, violating the norm of fairness). The replication of one central finding in the field of TMT speaks to the validity of the present study.

We want to note that we have not observed a relation of individual differences in everyday sadism and antisocial punishment in the absence of existential threat (i.e., in the control condition). This finding speaks to the notion that—like other motives (e.g., reciprocal behaviour; Schindler et al., 2013)—sadistic motives are not likely executed in a default mode but that specific situational boundary conditions are necessary for sadistic tendencies to be shown. It is plausible to assume that individuals not always show antisocial tendencies across situations but only when they are relevant and meaningful (e.g., to reestablish self-esteem after being threatened). As with other basic motivational orientations (such as prevention and promotion focus; Higgins, 1997, Pfattheicher & Sassenrath, 2014), situational boundary conditions are necessary to activate certain motivational orientations (cf. Higgins, 2012). It is reasonable to assume that this is also the case concerning sadistic tendencies. Yet future research can examine additional boundary conditions given the variation in dark personality traits (and antisocial punishment) across different societies or cultures (Grijalva & Newman, 2015, Herrmann et al., 2008).

Further explaining the absent relation of individual differences in sadism and antisocial punishment in the control condition, it is also possible that in a default mode, everyday sadists are as motivated as non-sadists to obtain monetary reward and avoid costly antisocial punishment. Also, given that our study included everyday sadists, it is plausible to assume that sadistic tendencies are not always and across situations applied by everyday sadists. That is, the motivation of maximizing benefits could be a driving force in everyday sadists and non-sadists, which leads to avoidance of antisocial punishment in a default mode.

The present work not only contributes to a better understanding of antisocial punishment but also contributes to research on everyday sadism. So far, sadism has not received much attention in the field of social psychology. This research shows a behavioural tendency (i.e., antisocial punishment) that corresponds to the main characteristics of sadism. That is, the present research's findings suggest that antisocial punishment appears to be a behavioural tendency executed by sadists. We have further revealed a boundary condition under which everyday sadists are likely to engage in harming others (i.e., when their self is existentially threatened). Thus, the present research contributes to the understanding of under what conditions everyday sadists may engage in evil behaviour and to the theory on everyday sadism.

We would like to acknowledge that our results may not be specific to existential threats to the self. In fact, other research has threatened the self differently (e.g., through a bad evaluation of one's essay; Bushman & Baumeister, 1998) and obtained strong aggressive responses following this threat to one's self. That is to say, future research could contribute to the question of whether the observed effects are

DOI: 10.1002/per

specific to existential threats. Given that we have argued that antisocial punishment may reflect a type of behaviour that allows the maintenance of self-esteem through aggressively dominating others, we would also expect that antisocial punishment is more likely to emerge when the self of everyday sadists is threatened in other ways.

At this point, we would also like to acknowledge that the present work remains silent regarding the relation of antisocial punishment and sadism to other antisocial traits, for instance, the dark triad of psychopathy, narcissism and Machiavellianism (e.g., Jones & Paulhus, 2012). In fact, one can conceptually distinguish sadism from the dark triad (Buckels et al., 2013; Chabrol, Meliolia, Van Leeuwena, Rodgers, & Goutaudiera, 2015). Psychopaths are empathy-lacking, impulsive individuals, ready to hurt others to serve their selfish goals, especially when being physically insulted (Jones & Paulhus, 2010). The core of narcissism is self-ascribed grandiosity and the ego-focus narcissists possess and their readiness to hurt others when being personally insulted (Bushman & Baumeister, 1998). Finally, individuals possessing a strong Machiavellian tendency are intensely focused on what will benefit them and are ready to exploit others. Empirically, Buckels et al. (2013) were able to document that narcissism and psychopathy were positively related to harming innocent others. However, only sadism was related to harming innocent others when harming others involved personal costs. Accordingly, given that antisocial punishment involves costs, we would predict that only sadism should be related to the costly behaviour of antisocial punishment, whereas psychopathy, narcissism and Machiavellianism should not predict this tendency. Future research could empirically clarify this point.

To conclude, the present work represents a new and promising approach to the study of antisocial punishment and its underlying forces and boundary conditions. It takes into account that behaviour in social dilemma situations is heavily influenced by the interplay of personal and situational factors. The results indicate the importance of taking antisocial personality variables (e.g., sadism) and situational boundary conditions (e.g., existential threat) as well as their interplay into account in order to reach a comprehensive understanding of individuals' behaviour in social dilemma situations.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

REFERENCES

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publications.

Arndt, J., Greenberg, J., & Cook, A. (2002). Mortality salience and the spreading activation of worldview-relevant constructs: Exploring the cognitive architecture of terror management. *Journal of Experimental Psychology: General*, *131*, 307–324.

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

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

Baumeister, R. F., Bushman, B. J., & Campbell, W. K. (2000). Self-esteem, narcissism, and aggression. Does violence result from low self-esteem or from threatened egotism? *Current Directions in Psychological Science*, *9*, 26–29.

Baumeister, R. F., & Campbell, W. K. (1999). The intrinsic appeal of evil: Sadism, sensational thrills, and threatened egotism. *Personality and Social Psychology Review*, *3*, 210–221.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, *103*, 5–33.

Baumeister, R. F., & Vohs, K. D. (2004). Four roots of evil. In A. G. Miller (Ed.), *The social psychology of good and evil* (pp. 85–101). New York: The Guilford Press.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*, 912–915.

Brown, R. P., & Zeigler-Hill, V. (2004). Narcissism and the non-equivalence of self-esteem measures: A matter of dominance? *Journal of Research in Personality*, *38*, 585–592.

Buckels, E. E., Jones, D. N., & Paulhus, D. L. (2013). Behavioral confirmation of everyday sadism. *Psychological Science*, *24*, 2201–2209.

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences*, *67*, 97–102.

Burke, B. L., Martens, A., & Faucher, E. H. (2010). Two decades of terror management theory: A meta-analysis of mortality salience research. *Personality and Social Psychology Review*, *14*, 155–195.

Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology*, *75*, 219–229.

Carré, J. M., & Mehta, P. H. (2011). Importance of considering testosterone-cortisol interactions in predicting human aggression and dominance. *Aggressive Behavior*, *37*, 1–3.

Chabrol, H., Melioli, T., Van Leeuwen, N., Rodgers, R., & Goutaudier, N. (2015). The Dark Tetrad: Identifying personality profiles in high-school students. *Personality and Individual Differences*, *83*, 97–101.

Chabrol, H., Van Leeuwen, N., Rodgers, R., & Séjourné, N. (2009). Contributions of psychopathic, narcissistic, Machiavellian, and sadistic personality traits to juvenile delinquency. *Personality and Individual Differences*, *47*, 734–739.

Cooke, D. J. (2001). Psychopathy, sadism and serial killing. In A. Raine & J. Sanmartín (Eds.), *Violence and psychopathy* (pp. 123–137). New York: Springer.

Crocker, J., Lee, S. J., & Park, L. E. (2004). The pursuit of self-esteem: Implications for good and evil. In A. G. Miller (Ed.), *The social psychology of good and evil* (pp. 271–302). New York: The Guilford Press.

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, *130*, 392–414.

Dietz, P. E., Hazelwood, R. R., & Warren, J. (1990). The sexually sadistic criminal and his offenses. *Journal of the American Academy of Psychiatry and the Law Online*, *18*, 163–178.

Dreber, A., & Rand, D. G. (2012). Retaliation and antisocial punishment are overlooked in many theoretical models as well as behavioral experiments. *Behavioral and Brain Sciences*, *35*, 24–24.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425*, 785–791.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*, 137–140.

Fein, S., & Spencer, S. J. (1997). Prejudice as self-image maintenance: Affirming the self through derogating others. *Journal of Personality and Social Psychology*, *73*, 31–44.

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

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*, 171–178.

DOI: 10.1002/per

Fowler, J. H., Johnson, T., & Smirnov, O. (2004). Egalitarian motive and altruistic punishment. *Nature*, *433*, E1–E2.

Fritsche, I., Jonas, E., & Fankhänel, T. (2008). The role of control motivation in mortality salience effects on ingroup support and defense. *Journal of Personality and Social Psychology*, *95*, 524–541.

Fromm, E. (1973). *The anatomy of human destructiveness*. New York: Holt, Rinehart & Winston.

Gächter, S., & Herrmann, B. (2009). Reciprocity, culture and human cooperation: Previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B*, *364*, 791–806.

Goette, L., Huffman, D., Meier, S., & Sutter, M. (2012). Competition between organizational groups: Its impact on altruistic and anti-social motivations. *Management Science*, *58*, 948–960.

Greenberg, J., Pyszczynski, T., Solomon, S., Rosenblatt, A., Veeder, M., Kirkland, S., & Lyon, D. (1990). Evidence for terror management theory II: The effects of mortality salience on reactions to those who threaten or bolster the cultural worldview. *Journal of Personality and Social Psychology*, *58*, 308–318.

Greenberg, J., Solomon, S., & Arndt, J. (2008). A basic but uniquely human motivation: Terror management. In J. Y. Shah & W. L. Gardner (Eds.), *Handbook of motivation science* (pp. 114–134). New York: Guilford Press.

Greitemeyer, T. (2015). Everyday sadism predicts violent video game preferences. *Personality and Individual Differences*, *75*, 19–23.

Grijalva, E., & Newman, D. A. (2015). Narcissism and counterproductive work behavior (CWB): Meta-analysis and consideration of collectivist culture, Big Five personality, and narcissism's facet structure. *Applied Psychology*, *64*, 93–126.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis*. New York: The Guilford Press.

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

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Higgins, E. T. (2012). *Beyond pleasure and pain: How motivation works*. Oxford: Oxford University Press.

Irwin, K., & Horne, C. (2013). A normative explanation of antisocial punishment. *Social Science Research*, *42*, 562–570.

Jonas, E., Martens, A., Niesta, D., Fritsche, I., Sullivan, D., & Greenberg, J. (2008). Focus theory of normative conduct and terror management theory: The interactive impact of mortality salience and norm salience on social judgment. *Journal of Personality and Social Psychology*, *95*, 1239–1251.

Jones, D. N., & Paulhus, D. L. (2010). Different provocations trigger aggression in narcissists and psychopaths. *Social Psychological and Personality Science*, *1*, 12–18.

Jones, D. N., & Paulhus, D. L. (2012). Differentiating the dark triad within the interpersonal circumplex. In L. M. Horowitz & S. N. Strack (Eds.), *Handbook of interpersonal theory and research* (pp. 249–267). Hoboken, NJ: Wiley & Sons.

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer Deutschen version der "Positive and Negative Affect Schedule" (PANAS) [studies with a German version of the "Positive and Negative Affect Schedule" (PANAS)]. *Diagnostica*, *42*, 139–156.

Leary, M. R. (1999). The social and psychological importance of self-esteem. In R. M. Kowalski & M. R. Leary (Eds.), *The social psychology of emotional and behavioral problems: Interfaces of social and clinical psychology* (pp. 197–221). Washington, DC: American Psychological Association.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. *32*, pp. 1–62). New York: Academic Press.

Leary, M. R., Cottrell, C. A., & Phillips, M. (2001). Deconfounding the effects of dominance and social acceptance on self-esteem. *Journal of Personality and Social Psychology*, *81*, 898–909.

McDonald, J. F., & Moffitt, R. A. (1980). The uses of Tobit analysis. *The Review of Economics and Statistics*, *62*, 318–321.

Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and Behavior*, *58*, 898–906.

Minson, J. A., & Monin, B. (2012). Do-gooder derogation disparaging morally motivated minorities to defuse anticipated reproach. *Social Psychological and Personality Science*, *3*, 200–207.

Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology*, *95*, 76–93.

O'Meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the short sadistic impulse scale (SSIS). *Psychological Assessment*, *23*, 523–531.

Ostrom, E. (1990). *Governing the commons*. Cambridge, MA: Cambridge University Press.

Parks, C. D., & Stone, A. B. (2010). The desire to expel unselfish members from the group. *Journal of Personality and Social Psychology*, *99*, 303–310.

Paulhus, D. L., & Jones, D. N. (in press). Measuring dark personalities. In G. J. Boyle, D. H. Saklofske & G. Matthews (Eds.), *Measures of personality and social psychological constructs*. San Diego, CA: Academic Press.

Pfattheicher, S., & Keller, J. (2013). Vigilant self-regulation and costly punishment in public goods situations. *European Journal of Personality*, *27*, 346–354.

Pfattheicher, S., & Keller, J. (2014). Towards a biopsychological understanding of costly punishment: The role of basal cortisol. *PLoS ONE*, *9*, e85691.

Pfattheicher, S., Landhäußer, A., & Keller, J. (2014). Individual differences in antisocial punishment in public goods situations: The interplay of cortisol with testosterone and dominance. *Journal of Behavioral Decision Making*, *27*, 340–348.

Pfattheicher, S., & Sassenrath, C. (2014). A regulatory focus perspective on eating behavior: How promotion and prevention focus relates to emotional, external, and restrained eating. *Frontiers in Psychology*, *5*, 1314.

Piven, J. S. (2003). Death, repression, narcissism, misogyny. *The Psychoanalytic Review*, *90*, 225–260.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*, 185–227.

Pyszczynski, T., Greenberg, J., Solomon, S., Arndt, J., & Schimel, J. (2004). Why do people need self-esteem? A theoretical and empirical review. *Psychological Bulletin*, *130*, 435–468.

Reidy, D. E., Zeichner, A., & Seibert, L. A. (2011). Unprovoked aggression: Effects of psychopathic traits and sadism. *Journal of Personality*, *79*, 75–100.

Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T., & Lyon, D. (1989). Evidence for terror management theory: I. The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, *57*, 681–690.

Schindler, S., Reinhard, M.-A., & Stahlberg, D. (2013). Tit for tat in the face of death: The effect of mortality salience on reciprocal behavior. *Journal of Experimental Social Psychology*, *49*, 87–92.

Singer, T., & Steinbeis, N. (2009). Differential roles of fairness-and compassion-based motivations for cooperation, defection, and punishment. *Annals of the New York Academy of Sciences*, *1167*, 41–50.

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

Tabibnia, G., Satpute, A., & Lieberman, M. (2008). The sunny side of fairness. *Psychological Science*, *19*, 339–347.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.

Twenge, J. M., Baumeister, R. F., Tice, D. M., & Stucke, T. S. (2001). If you can't join them, beat them: Effects of social

DOI: 10.1002/per

exclusion on aggressive behavior. *Journal of Personality and Social Psychology*, *81*, 1058–1069.

Twenge, J. M., & Campbell, W. K. (2003). "Isn't it fun to get the respect that we're going to deserve?" Narcissism, social rejection, and aggression. *Personality and Social Psychology Bulletin*, *29*, 261–272.

van Lange, P. A. M., Balliet, D. P., Parks, C. D., & van Vugt, M. (2014). *Social dilemmas: Understanding human cooperation*. Oxford: Oxford University Press.

van Lange, P. A. M., Joireman, J., Parks, C. D., & van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, *120*, 125–141.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

DOI: 10.1002/per