Organizational Behavior and Human Decision Processes 149 (2018) 59–72

Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

From spontaneous cooperation to spontaneous punishment – Distinguishing the underlying motives driving spontaneous behavior in first and second order public good games

Dorothee Mischkowski[a,b,*], Andreas Glöckner[b,c], Peter Lewisch[d]

[a] University of Göttingen, Gosslerstrasse 14, D-37073 Göttingen, Germany
[b] University of Hagen, Universitätsstrasse 27, D-58097 Hagen, Germany
[c] Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn, Germany
[d] University of Vienna, Schenkenstraße 4, A-1010 Vienna, Austria

ARTICLE INFO

Keywords:
Spontaneous cooperation
Spontaneous punishment
Decision time
Second-party punishment
Social Value Orientation
Above-average contributors
Negative affect

ABSTRACT

Recent findings indicate that at least some individuals use prosocial options by default in social dilemmas, known as 'spontaneous cooperation'. In two studies, we test whether this effect generalizes to second order public goods in the form of punishment behavior in one-shot and iterated public good games and investigate the underlying motivations. In line with spontaneous cooperation, punishment decreases with increasing decision time. Negative affect moderates this spontaneous punishment effect in one-shot public good games, in that punishment decisions are made more quickly by persons who are more upset about the contribution behavior of their group members. Unlike spontaneous cooperation, spontaneous punishment is not directly influenced by dispositional pro-sociality but by situationally above-average contributions. An overall analysis indicates a three-way interaction in that the spontaneous punishment effect is mainly valid for above-average, highly upset contributors. Hence, our results highlight the uniqueness of spontaneous punishment as being, in contrast to spontaneous cooperation, an affect-driven phenomenon of above-average contributors.

1. Introduction

Whether humans are intrinsically good or evil is a key question that is discussed in many academic disciplines such as philosophy, economics, and psychology, to name but a few. Research examining situations in which maximizing individual gains conflicts with the community's welfare (e.g., in social dilemmas) has revealed insights into the psychological mechanisms of cooperation behavior (e.g., Hardin, 1968). Recent research that went beyond a mere analysis of choices by including process measures such as decision times and eye-tracking has made a significant contribution to our knowledge concerning the underlying cognitive processes of prosocial behavior (De Dreu et al., 2010; Fiedler, Glöckner, Nicklisch, & Dickert, 2013; Rand, Greene, & Nowak, 2012; Rubinstein, 2007; Fischbacher, Hertwig, & Bruhin, 2013; Lotito et al., 2013). In an influential publication, Rand et al. (2012) suggested that cooperation behavior is more spontaneous than defection and accompanied by shorter decision times and can even be enhanced by manipulation (e.g., time pressure) to impose more intuitive processing. This indicates that the answer to the introductory question could be a positive one, in that humans might indeed be "of a good kind" and that pro-sociality can be promoted even further.

This so-called 'spontaneous cooperation effect' has inspired much subsequent research. On the one hand, many replication studies focused on the causality of the effect by manipulating intuitive processing. Several of those studies successfully replicated the effect (e.g., Rand et al., 2014; Lotz, 2015; Protzko, Ouimette, & Schooler, 2016), whereas others did not (e.g., Tinghög et al., 2013; Verkoeijen & Bouwmeester, 2014). Recently, Rand (2016) published a meta-analysis that considers different experimental manipulations of intuitive processing and shows a strong positive relationship between intuition and cooperation. Importantly, further analyses revealed that there was no statistical evidence of publication bias in the data set. Simultaneously, a many-labs, pre-registered replication project was launched to investigate the effect of time pressure / time delay as one specific manipulation of processing mode more broadly, while preventing publication biases in advance (see Open Science Framework https://osf.io/scu2f/; Bouwmeester et al., 2017). Aggregating over all 21 labs and 3596 participants, the spontaneous cooperation effect was not replicated. Still, the results

* Corresponding author at: FernUniversität in Hagen, Department of Psychology, Universitätsstr. 27, D-58097 Hagen, Germany.
E-mail address: dorothee.mischkowski@fernuni-hagen.de (D. Mischkowski).

https://doi.org/10.1016/j.obhdp.2018.07.001
Received 4 August 2016; Received in revised form 4 June 2018; Accepted 10 July 2018

painted a heterogeneous picture in that a small-sized effect reappeared when subjects who did not comply with the time pressure/time delay manipulation were excluded – thus opening the discussion for potential selection effects (Bouwmeester et al., 2017; Rand, 2017).

In contrast, several studies applied the methodology of the first, correlative study from the original publication by Rand et al. (2012) by measuring rather than manipulating decision times (e.g., Nielsen, Tyran, & Wengström, 2014). In this paradigm, the negative correlation between decision time and cooperation was repeatedly replicated (e.g., Lotito et al., 2013; Cappelen, Nielsen, Tungodden, Tyran, & Wengström, 2016). Further studies, however, showed that the effect is highly volatile and influenced by multiple moderators, which relate mainly to personality characteristics such as Social Value Orientation (SVO, Mischkowski & Glöckner, 2016), Honesty-Humility (Kieslich & Hilbig, 2014; Mischkowski & Glöckner, 2016) and trust in the cooperativeness of daily-life interaction partners (Rand et al., 2012). These conditional effects indicate that spontaneous cooperation is the default only for persons with a prosocial, honest, or trusting attitude.

Researchers have questioned whether decision time measurement within the correlational paradigm investigating the spontaneous cooperation effect allows for conclusions concerning dual-process models – i.e., whether decisions are made in a more intuitive or deliberate manner (Krajbich, Bartling, Hare, & Fehr, 2015). Decision time measurement is generally evaluated as an important and valuable indicator for the identification of decision processes (e.g., Glöckner & Betsch, 2008a; Fiedler et al., 2013; Heck & Erdfelder, 2017; Spiliopoulos & Ortmann, 2018) that is inherently free of selection bias. However, it is complicated by the fact that it appears to be driven by multiple factors. Deliberation, per definition, should require more time than intuition. In addition, prior evidence in various domains has shown that decision conflict, defined as the subjective discriminability of choice options (i.e., difference in utility between choice options or, more broadly, strength-of-preference or differentiation in the phenomenological field; Cartwright & Festinger, 1943), also determines decision time (e.g., Festinger, 1943; Birnbaum & Jou, 1990). Specifically, the more similar decision options are evaluated to be due to similar aspects or cues, the greater the increase in conflict and, hence, decision time (e.g., Fiedler et al., 2013; Glöckner & Betsch, 2012). In the domain of cooperation behavior, Evans, Dillon, and Rand (2015) also show that decision times reflect the extent of decision conflict (see also Krajbich et al., 2015). Particularly, Evans et al. (2015) found that extreme (i.e., highly cooperative as well as highly defective) decisions are associated with quicker decision times than intermediate decisions, thus providing evidence for an inverse u-shaped pattern of the relation between cooperation behavior and decision time. Decision times not only increased with decision conflict but also mediated the relation between levels of conflict and decision extremity. Finally, manipulated decision time led to higher cooperation behavior but did not influence decision extremity and feelings of conflict. The authors conclude that only manipulated decision time should be used to interpret the degree of intuitiveness as measured decision time always incorporates the level of decision conflict.

Taken these findings together, short decision times might be a reflection of an individual's default manner of decision making in situations with low decision conflict, which are associated with a reduced necessity to invest cognitive resources to inhibit or alter default responses (see also Evans, 2008; Horstmann, Ahlgrimm, & Glöckner, 2009). Low decision conflict can be based on both, individual (i.e., the person generally tends to behave in that manner) as well as situational factors (e.g., all cues of the environment speak for one of the options). In line with this argument, Nishi, Christakis, Evans, O'Malley, and Rand (2016) show that the relationship between measured (endogenous) decision time and cooperation behavior in repeated games depends on both the social environment (i.e., on the cooperation level of the previous partner) and the player's own level of cooperativeness. These findings indicate that individuals' decision behavior is shaped not only by one's personality characteristics (i.e., own cooperativeness) but also by the environment.

This paper approaches the spontaneous cooperation effect from a different perspective, namely by investigating whether it generalizes to punishment decisions (see paragraph below). That is, we tested whether punishment investments involve similar properties concerning cognitive processes as investments in the first order public good (i.e., cooperation behavior). Additionally, we investigated the motivational forces accounting for this (assumed) effect. In doing so, we assessed the possibilities that spontaneous punishment could be based on a revenge-oriented, affect-driven behavior as well as that it could be dependent on traits focusing on equality and fairness, i.e., SVO, in line with spontaneous cooperation for prosocials.

1.1. Punishment behavior in social dilemmas

Punishment in social dilemmas consists of a costly sanction mechanism for norm violators (i.e., distribution norms; Fehr & Fischbacher, 2004; see also Yamagishi, 1986). It constitutes a second order public good in social dilemmas, in that it follows the same incentive structure as cooperation behavior, representing a first order public good. In first order public goods, an individual is confronted with the decision of whether to contribute (monetary) resources to a common pool that benefits the whole group to a larger extent than the individual – with no knowledge of the group members' simultaneous behavior. In contrast, punishment requests an individual's monetary investment to reduce the income of another person, thereby providing the opportunity to punish previous non-cooperative behavior and generate equality among all group members. In anonymous repeated public good situations, individuals usually show substantial cooperation behavior in the beginning that, however, deteriorates over time (Fehr & Gächter, 2002; Egas & Riedl, 2008). The option to punish was introduced as an effective instrument with which to maintain high levels of cooperation (Fehr & Gächter, 2002, Yamagishi, 1986, 1988). Generally, punishment can be conducted by persons that are part of the group (second-party punishment), third-parties who are outside observers of the group setting (altruistic punishment), or fixed punishment mechanisms (Yamagishi, 1986). In the current paper, we consider only the two former cases of punishment. In these types of punishment, after having seen the contributions of other players, participants decide whether or not to invest money to reduce the profit of other players who, for example, have been free-riding on the contributions made by other persons. Thus, punishment represents a second order public good in that "everybody in the group would be better off if free riding is deterred and high levels of cooperation are sustained, but nobody has an incentive to punish the free riders" (Fehr & Gächter, 2002, p. 137) and would preferably leave the costs of punishment to other group members.

To explain in psychological terms why people inflict considerable punishment even given the (second order) public good structure, Fehr and Gächter (2002) suggest that emotions drive altruistic punishment. In one-shot settings, the retributive character of punishing free-riders might be conveyed by anger. In repeated settings, this might be entangled with a future-oriented motivation to deter further free-riding.[1] Importantly, negative affect related to punishment decisions can be further specified. Specifically, researchers have argued that particularly

[1] Empirical results testing the differences between both punishment designs show that second-party punishment is usually conducted more often and more strongly in comparison to third-party punishment (Fehr & Fischbacher, 2004). Several articles shed light on the differences of elicited emotions dependent on whether second- or third-party punishment was conducted (e.g., Fehr & Fischbacher, 2004; Fetchenhauer & Huang, 2004). Consensus is that self-experienced injustice leads to stronger anger whereas altruistic punishment is more closely related to moral indignation (Camerer, 2003).