From Spontaneous Cooperation to Spontaneous Punishment: Distinguishing the Underlying Motives Driving Spontaneous Behavior in First and Second Order Public Goods

Peter Lewisch, Dorothee Mischkowski, Andreas Glöckner
Rand, Greene, & Nowak (2012) see cooperation behavior as the first and intuitive response in social dilemmas (e.g., Public Good Games)
... is driven by prosocials

Mischkowski & Glöckner (2016) show that there is a spontaneous cooperation effect for prosocials, but not for proself persons.
The influence of social value orientation in strategic interactions

SVO represents a continuous measure of social preferences, “defined in terms of the weights people assign to their own and others’ outcomes in situations of interdependence” (Balliet, Parks, & Joireman, 2009, p.533)

\[ U = w_1 \times u \text{ (own payoff)} + w_2 \times u \text{ (other’s payoff)} \]
Further moderators driving the spontaneous cooperation effect

Trait dependency:
- Honesty-Humility as one of the six basic trait factors in the HEXACO model (Kieslich & Hilbig, 2015; Mischkowski & Glöckner, 2016)
- Cooperativeness of daily life interaction partners (Rand et al., 2012)

Situational (state) dependency:
- Experience in social dilemma games reduces spontaneous cooperation in one-shot dilemma games (Rand et al., 2012)
Altruistic Punishment

Fehr & Gächter (2002) show that third-party punishment / altruistic punishment (AP) is suited and used to enhance cooperation in repeated games.

Participants invest their own money to reduce the outcome of other players.

AP represents a second order public good as 'everybody in the group will be better off if free riding is deterred, but nobody has an incentive to punish the free riders' (Fehr & Gächter, 2002, p.137).
Underlying Factors of Altruistic Punishment

Punishment is mainly driven by above-average cooperators
(Fehr & Gächter, 2002; Fehr & Fischbacher, 2004; Fehr, Fischbacher, & Gächter, 2002)

„Emotions are an important proximate factor behind altruistic punishment“
(Fehr & Gächter, 2002, p.139)

(Russel & Feldman Barrett, 1999)
Research Question

Can we expand the spontaneous cooperation effect to punishment behavior?

→ Does punishment imply the same mechanisms as cooperation behavior (resulting in the same behavioral pattern)?

→ If so, is spontaneous punishment a revenge-oriented, affect-driven behavior or rather trait-dependent on SVO to restore justice?
Hypotheses

1. Punishment decisions are conducted quicker than non-punishment decisions

2. Analogue to spontaneous cooperation, spontaneous punishment is driven by prosocials. For proselfs, no difference in punishment behavior is expected over time

3. Spontaneous punishment goes in line with negative affect
Method

Two studies ($N_{\text{total}} = 277$) in which groups of 4 players engage in a Public Goods Game (PGG, MPCR = .5)

Study 1:
- Iterative PGG: two rounds, no stranger matching for direct deterrence
- $N = 123$ (65 female), mean age = 22.7 years ($SD = 5.0$)

Study 2:
- One-shot PGG, $N = 154$, 78 female,
- Mean age = 23.29 years ($SD = 5.2$)
Method

- Second-party punishment (*Cost to Impact Ratio = 1:4*)
  - Measurement of the punishment decision time for the binary decision („Do you want to punish at least one of your team members?“)
  - Specification of the amount of punishment on a subsequent screen

- Affect was captured after the PGG (via the PANAS)

- In order to separate the measurement of social preferences from cooperation behavior, SVO was measured up to 12 hours before the lab session (SVO Slider Measure, Murphy, Ackermann, Handgraaf, 2011)
Results: Spontaneous Punishment (H1)

- Overall, there is no spontaneous punishment for the binary punishment decision (logistic regression: odds ratio = .42, $z = -1.14$, $p = .255$)

- However, when differentiating the amount of punishment (no punishment = 0 Cent Investment), a continuous decrease of punishment investments turn visible (OLS regression: $\beta = - .19$, $t(275) = -3.16$, $p = .002$)
Results: Spontaneous Punishment & SVO (H2)

• Spontaneous Punishment – unlike spontaneous cooperation – is not driven by prosocials, neither for the binary punishment decision (odds ratio = .98, \( z = -.30, p = .76 \)), nor for the continuous punishment investment (\( \beta = -.03, t(260)= -.46, p = .647 \)).

• Interestingly, there is even no main effect of SVO on Punishment (\( r = -.07, p = .23 \)) nor for Inequality Aversion - as a facet of prosociality - and punishment behavior (\( r = -.05, p = .625 \))

→ Alone in study1, there is a tendency for people who are inequality averse to punish more frequently than joint gain maximizer (JGM) (\( r = .25, p = .075 \))
Results: Spontaneous Punishment & Cooperation Behavior

- In line with Fehr & Gächter (2002) who identified above average cooperation as one source of punishment, spontaneous punishment decisions are driven by above average contributors (binary pun.: odds ratio = 0.03, z = -1.92, p = .055; pun. investments: $\beta = -0.14$, $t(273) = -2.33$, $p = .02$)
- However, contributions are only moderately related to SVO ($r = .19$, $p = .02$)
Results: Spontaneous Punishment & Negative Affect (H3)

- Spontaneous punishment decisions are driven by negative affect, only marginally significant for binary punishment decisions (odds ratio = .63, z = -1.70, p = .090), but highly significant for the continuous punishment investment ($\beta = -.26$, $t(272) = -4.20$, $p < .001$).
Conclusion

• Similar to a spontaneous cooperation effect, we find that the invested resources to punish decrease with increasing decision time. There is no main effect of spontaneous punishment behavior for the binary punishment decision, though.

• In contrast to spontaneous cooperation behavior that is driven by prosocials, spontaneous punishment is an emotional act, driven by negative affect.

• Even though dispositional social preferences are not moderating spontaneous punishment, situational above average cooperation behavior does promote spontaneous punishment.

• Punishment as a second order public good therefore seems to follow a different motivation than cooperation behavior, even though the mere behavior follows a similar pattern.
Thank you very much for your attention!