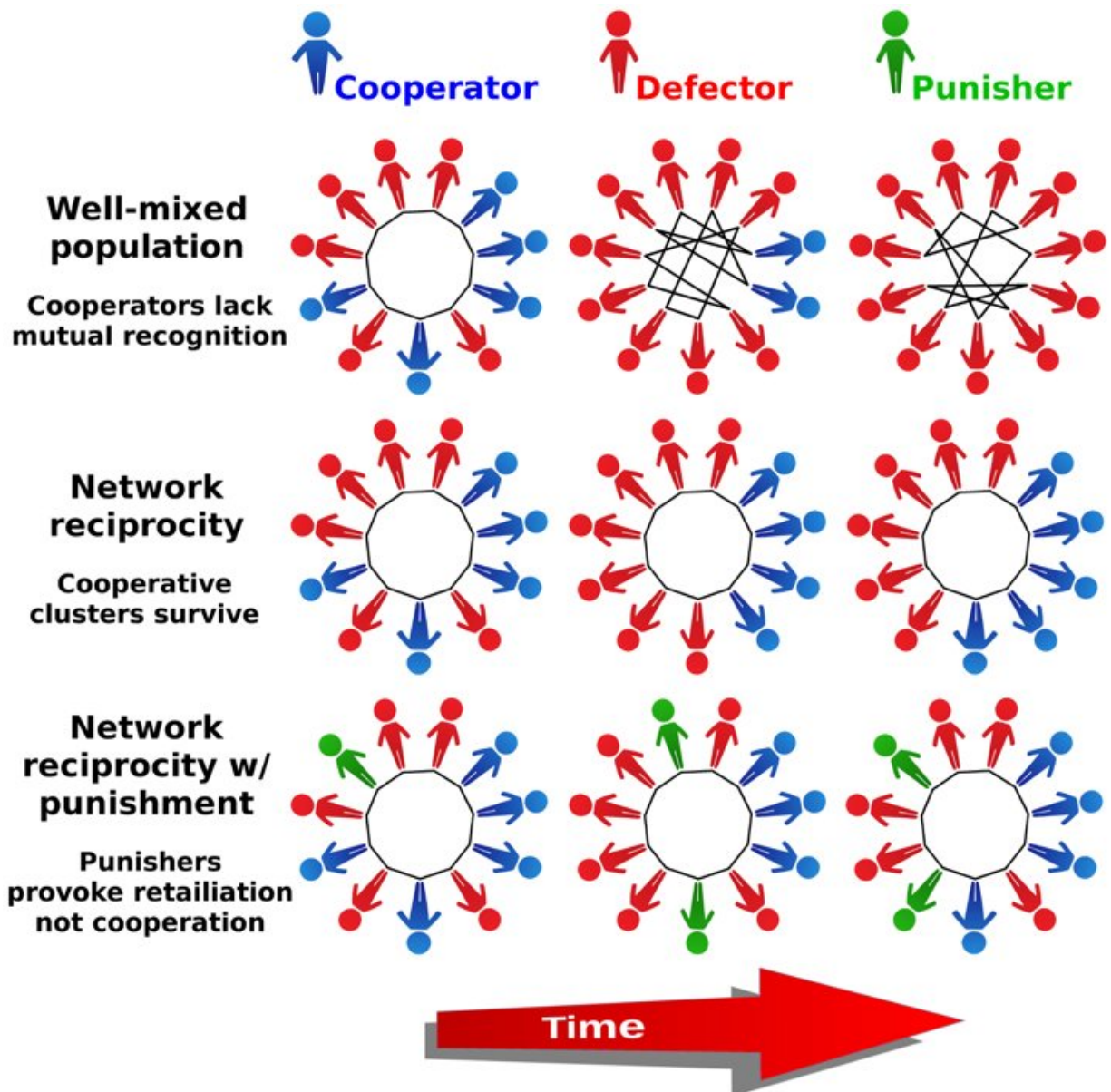


Is punishment as effective as we think?

December 27 2017



In the first group (Well-mixed) where the opponents were reshuffled each round, defectors prevailed over the course of 50 rounds. In the second group (Network

reciprocity) where the opponents remained the same for the 50 rounds enabling them to identify cooperative neighbors, the cooperative cluster survived. In the third group (Network reciprocity with punishment), the option to punish opponents failed to boost cooperation. Credit: Hokkaido University

A game to study human behavior has shown punishment is an ineffective means for promoting cooperation among players. The result has implications for understanding how cooperation has evolved to have a formative role in human societies.

Human societies maintain their stability by forming cooperative partnerships. But, cooperation often comes at a cost. For example, a person taking time to raise the alarm in order to alert other members of a group to impending danger could be losing valuable time to save oneself. It is unclear why natural selection favors cooperativeness among individuals who are inherently selfish.

In theoretical studies, punishment is often seen as a means to coerce people into being more cooperative. To examine such theory, a team of international researchers led by Marko Jusup of Hokkaido University in Japan and Zhen Wang of Northwestern Polytechnical University in China has conducted a "social dilemma experiment." The team investigated if providing punishment as an option helps improve the overall level of cooperation in an unchanging network of individuals.

They used a version of the commonly employed "prisoner's dilemma" [game](#). Two hundred and twenty-five students in China were organized into three trial groups and played 50 rounds each of the game.

In group one, every student played with two opponents which changed every round. The students could choose between "cooperate" or "defect",

and [points](#) were given based on the combined choices made. If a student and the two opponents chose "defect," the student gained zero points. If they all chose "cooperate," the student gained four points. If only a student chose to defect while the other two chose to cooperate, the gain for the [student](#) was eight points.

The second group was similar to the first one in every aspect except that the people playing the game with each other remained the same for the duration of the 50 rounds, enabling them to learn each other's characteristics.

In the third group, players also remained the same. However, a new option, "punish," was introduced. Choosing punishment led to a small reduction in points for the punisher and a larger reduction of points for the punishees.

At the end of the game, overall points were counted and the students were given monetary compensation based on the number of points won.

The expectation is that, as individuals play more with the same opponents over several rounds, they see the benefit of cooperating in order to gain more points. Introducing punishment as an option is basically saying: if you don't cooperate with me, I'll punish you. In theory, it is expected that applying this option would lead to more cooperation.

The researchers found that players in the constantly changing groups cooperated much less (4%) than those in the static groups (38%), where they were able to establish which players were willing to cooperate and thus gain a larger average financial payoff for all involved.

Surprisingly, however, adding punishment as an option did not improve the level of cooperation (37%). The final financial payoffs in this trial

group were also, on average, significantly less than those gained by players in the static group. Interestingly, less defection was seen in the punishment group when compared to the static [group](#); some players replaced defection with punishment.

"While the implied message when punishing someone is 'I want you to be cooperative,' the immediate effect is more consistent with the message 'I want to hurt you,'" write the researchers in their study published in the journal *Proceedings of the National Academy of Sciences*.

Punishment seems to have an overall demoralizing effect, as individuals who get punished on multiple occasions may see a good chunk of their total payoff vanish in a short period of time, explain the researchers. This could lead players to lose interest in the game and play the remaining rounds with less of a rational strategy. The availability of punishment as an option also seems to reduce the incentive to choose cooperation over competition.

Why, then, is [punishment](#) so pervasive in [human societies](#)? "It could be that human brains are hardwired to derive pleasure from punishing competitors," says Jusup. "However, it is more likely that, in real life, a dominant side has the ability to punish without provoking retaliation," adds Wang.

Although the study provides valuable insights into how [cooperation](#) arises in human society, the team advises it would be unwise to extrapolate the implications of their study far beyond the experimental setting.

More information: Xuelong Li et al. Punishment diminishes the benefits of network reciprocity in social dilemma experiments, *Proceedings of the National Academy of Sciences* (2017). [DOI](#):

[10.1073/pnas.1707505115](https://doi.org/10.1073/pnas.1707505115)

Provided by Hokkaido University

Citation: Is punishment as effective as we think? (2017, December 27) retrieved 5 October 2025 from <https://phys.org/news/2017-12-effective.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.