

# Designing public institutions that foster cooperation

June 14 2021

---



Credit: CC0 Public Domain

Humans often cooperate, but ample research has shown that they're conditionally cooperative; that is, they are far more likely to cooperate with those who they consider "good."

In large societies, however, people don't always know the reputations of the people with whom they interact. That's where [reputation](#) monitoring systems—such as the star ratings for eBay sellers or the scores assigned by credit bureaus—come into play, helping guide people's decisions about whether or not they want to help or interact with another person.

In a new paper in the journal *Nature Communications*, a team from Penn uses mathematical modeling to study how public institutions of reputation monitoring can foster cooperation and also encourage participants to adhere to its assessments instead of relying on their own subjective judgments of each others' reputations.

"We show how to construct institutions of public monitoring that foster cooperation, regardless of the social norm of moral judgment," says Joshua Plotkin, a professor in the Department of Biology in Penn's School of Arts & Sciences who coauthored the paper with postdoctoral fellows Arunas Radvilavicius and Taylor Kessinger. "And then adherence to the public institution will naturally spread."

The work explores the concept known as indirect reciprocity. Unlike direct reciprocity, in which two people may take turns helping one another, indirect reciprocity depends on a shared moral system.

"Under the theory of indirect reciprocity, if I encounter someone who is known to be good, then I'll probably cooperate with them, even without any tangible benefit to myself," Plotkin says. "By doing this I gain something intangible—[social capital](#), or reputation—that is potentially valuable down the line. I'll be seen as a good person, and a third party may later repay my kindness. But if I defect against that good person, then I'll likely end up with a bad reputation, and I won't benefit from anyone else's help in the future."

Different social norms vary in how they assign moral reputations to

individuals based on their actions. Stemming from [game theory](#), one classic social norm is called "stern judging," in which cooperating with someone good earns you a good reputation, but cooperating with someone bad earns you a bad reputation. Another is "simple standing," "a more forgiving norm," Plotkin says, in which cooperating with someone bad also gains you a good reputation.

When studying how social norms might foster cooperation, however, prior studies assumed that everyone knows each other's reputations and that those views are all consistent. In the [real world](#), of course, people can make individual decisions about others' reputations. And when these views are inconsistent and opinions differ, "it can lead to a collapse of cooperation," says Plotkin.

One way to solve this is to have an institution offer a public assessment of each member's reputation. In the current work, the researchers aimed to test what features of such an institution will lead to the highest levels of cooperation and when individuals will adhere to the public broadcast.

They considered a scenario in which individuals could choose whether to make decisions based on their own perceptions of others' reputations, or choose to rely on the assessments of the designated public institution.

"You can imagine a simple institution consisting of just two observers, who compare their observations and come up with a consensus view of reputations to broadcast publicly," Plotkin says.

By varying the number of observers and the strictness with which they form their consensus views, the researchers found they could always get cooperation to flourish in their models, no matter which social norm was present—simple standing, stern judging, or others.

What's more, individuals evolved to adhere to the institution's

assessments.

"Even if only a few people in the population adhere to the institution's judgements to begin with," Plotkin says, "those individuals will be better off. And so institutional adherence will tend to spread by social contagion. So there's a nice sense in which we can specify institutions that foster cooperation and then get adherence for free."

In follow-up work, Plotkin and colleagues hope to probe what happens to cooperation and adherence under different scenarios. What happens when individuals must pay a "tax" to support a public monitoring system? Can such an [institution](#) resist corruption, or avoid bias? And what happens when a variety of social norms exist in a population? Such variables could bring the team's work closer to applications in human society.

"Unlike other theories of cooperation, which make sense for simple organisms such as bacteria," Plotkin says, "this study explores an explanation for [cooperation](#) that is compelling in human societies, where reputations are carefully monitored and valued."

**More information:** Arunas L. Radzvilavicius et al, Adherence to public institutions that foster cooperation, *Nature Communications* (2021). [DOI: 10.1038/s41467-021-23783-9](https://doi.org/10.1038/s41467-021-23783-9)

Provided by University of Pennsylvania

Citation: Designing public institutions that foster cooperation (2021, June 14) retrieved 26 April 2024 from <https://phys.org/news/2021-06-foster-cooperation.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.