

City University of New York (CUNY)

CUNY Academic Works

Publications and Research

CUNY Graduate Center

2022

Theory of Collective Action

Rohit J. Parikh
CUNY Graduate Center

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/gc_pubs/723

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).
Contact: AcademicWorks@cuny.edu

Theory of Collective Action

Rohit Parikh

Philog Seminar

October 2, 2022

From Ostrom, Elinor. "A behavioral approach to the rational choice theory of collective action: Presidential address, American Political Science Association, 1997." *American political science review* 92.1 (1998): 1-22.

Let me start with a provocative statement. You would not be reading this article if it were not for some of our ancestors learning how to undertake collective action to solve social dilemmas.

The theory of collective action is the central subject of political science.

We have not yet developed a behavioral theory of collective action based on models of the individual consistent with empirical evidence about how individuals make decisions in social-dilemma situations

Elinor Ostrom

Garrett Hardin's (1968) evocative paper, "The Tragedy of the Commons" is based on an assumption that rational individuals are helplessly trapped in social dilemmas from which they cannot extract themselves without inducement or sanctions applied from the outside. Many policies based on this assumption have been subject to major failure and have exacerbated the very problems they were intended to ameliorate.

Hardin, Garrett. "The tragedy of the commons." *Green Planet Blues*, Routledge, 2019. 41-49.

See also, Birbal, *Sau Syane et mat*

Social dilemmas occur whenever individuals in interdependent situations face choices in which the maximization of short-term self-interest yields outcomes leaving all participants worse off than feasible alternatives.

In a public-good dilemma, for example, all those who would benefit from the provision of a public good—such as pollution control, radio broadcasts, or weather forecasting—find it costly to contribute and would prefer others to pay for the good instead.

Prisoner's Dilemma

	C	D
C	4, 4	1, 5
D	5, 1	2, 2

(C,C) is better for both but D is dominant strategy for both and hence it is (D,D) which is the unique Nash equilibrium.

N person prisoner's dilemma

Note: N players choose between cooperating (C) or not cooperating (-C). When individuals cooperate, their payoffs are always lower than the j individuals who do not cooperate. The predicted outcome is that no one will cooperate and all players will receive X benefits. The temptation (T) not to cooperate is the increase in benefit any cooperator would receive for switching to not cooperating. If all cooperate, they all receive $G-X$ more benefits than if all do not cooperate.

If everyone contributes, they get a net positive benefit (G). However, everyone faces a temptation (I) to shift from the set of contributors to the set of those who do not contribute.

The **theoretical prediction** is that everyone will shift and that no one will contribute. If this happens, then the outcome will be at the intercept. The difference between the predicted outcome and everyone contributing is $G - X$. Since the less-valued payoff is at a Nash equilibrium, no one is independently motivated to change his or her choice, given the choices of other participants.

These situations are **dilemmas** because at least one outcome exists that yields greater advantage for all participants. Thus, a **Pareto-superior** alternative exists, but rational participants making isolated choices are not predicted to realize this outcome.

A conflict is posed between individual and group rationality. The problem of collective action raised by social dilemmas is to find a way to avoid *Pareto-inferior* equilibria and to move closer to the optimum. Those who find ways to coordinate strategies in some fashion receive a "cooperators' dividend" equal to the difference between the predicted outcome and the outcome achieved

Stag hunt game

	Hare	Stag
Hare	2, 2	2, 0
Stag	0, 2	4, 4

It is better to play stag **if you can trust** the other party.

Hunting hare is safer, you get 2 regardless of what the other does, but has a lower reward.

Centipede Game

An Extensive Form representation of a four-stage centipede game, which ends after four rounds with the money being split. Passing the coins across the table is represented by a move of R (going across the row of the lattice, sometimes also represented by A for across) and pocketing the coins is a move D (down the lattice). The numbers 1 and 2 along the top of the diagram show the alternating decision-maker between two players denoted here as 1 and 2, and the numbers at the bottom of each branch show the payoff for players 1 and 2 respectively.

The theoretical prediction, using backward induction and common knowledge of rationality, is that player 1 will choose *D* at the very first move, even though better payoffs are available later on.

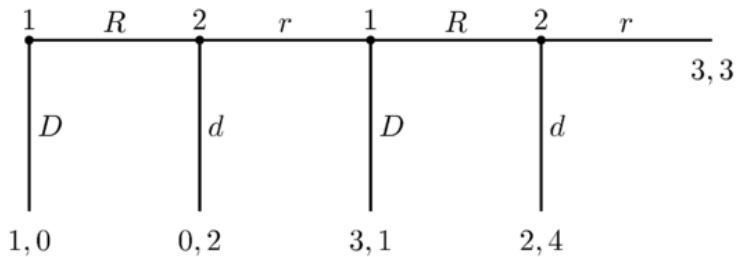


Figure: The centipede

Several studies have demonstrated that the Nash equilibrium (and likewise, subgame perfect equilibrium) play is rarely observed. Instead, subjects regularly show partial cooperation, playing "R" (or "r") for several moves before eventually choosing "D" (or "d"). It is also rare for subjects to cooperate through the whole game.

For examples see McKelvey and Palfrey (1992) and Nagel and Tang (1998). As in many other game theoretic experiments, scholars have investigated the effect of increasing the stakes. As with other games, for instance the ultimatum game, as the stakes increase the play approaches (but does not reach) Nash equilibrium play

She (Ostrom) summarizes four consistently replicated findings that directly challenge the general fit between behavior **observed** in social-dilemma experiments and the **predictions** of noncooperative game theory using complete rationality and complete information for one-shot and finitely repeated social dilemmas.

In other words, in practice people are more cooperative than theory predicts.

The four general findings are as follows.

1. High levels of initial cooperation are found in most types of social dilemmas, but the levels are consistently less than optimal.
2. Behavior is not consistent with backward induction in finitely repeated social dilemmas.
3. Nash equilibrium strategies are not good predictors at the individual level.
4. Individuals do not learn Nash equilibrium strategies in repeated social dilemmas.

TWO INTERNAL WAYS OUT OF SOCIAL DILEMMAS

Communication and Collective Action

From this theoretical perspective, face-to-face communication should make no difference in the outcomes achieved in social dilemmas. Yet, consistent, strong, and replicable findings are that substantial increases in the levels of cooperation are achieved when individuals are allowed to communicate face to face.

Sally (1995) finds that opportunities for face-to-face communication in one shot experiments significantly raise the cooperation rate, on average, by more than 45 percentage points.

When subjects are allowed to talk before each decision round in repeated experiments, they achieve 40 percentage points more on average than in repeated games without communication. No other variable has as strong and consistent an effect on results as face-to face communication.

Changing the rules of a game or using scarce resources to punish those who do not cooperate or keep agreements are usually not considered **viable options** for participants in social dilemmas, since these actions (only) create public goods.

Yet participants in many field settings and experiments do exactly this. Extensive research on how individuals have governed and managed common-pool resources has documented the incredible diversity of rules designed and enforced by participants themselves to change the structure of underlying social-dilemma situations.

The particular rules adopted by participants vary radically to reflect local circumstances and the cultural repertoire of acceptable and known rules used generally in a region. Nevertheless, general design principles characterize successfully self-organized, sustainable, local, regional, and international regimes (E. Ostrom 1990). Most robust and long-lasting common-pool regimes involve clear mechanisms for monitoring rule conformance and graduated sanctions for enforcing compliance

In field settings, innovation in rules usually occurs in a continuous trial-and-error process until a rule system is evolved that participants consider yields substantial net benefits. Given the complexity of the physical world that individuals frequently confront, they are rarely ever able to "get the rules right" on the first or second try (E. Ostrom 1990). In highly unpredictable environments, a long period of trial and error is needed before individuals can find rules that generate substantial positive net returns over a sufficiently long time horizon.

Nonviolent conflict may be a regular feature of successful institutions when arenas exist to process conflict cases regularly and, at times, to innovate new rules to cope with conflict more effectively.

In any population of individuals, one is likely to find some who use one of three reciprocity norms when they confront a repeated social dilemma.

- ▶ 1. Always cooperate first; stop cooperating if others do not reciprocate; punish noncooperators if feasible.
- ▶ 2. Cooperate immediately only if one judges others to be trustworthy; stop cooperating if others do not reciprocate; punish noncooperators if feasible.
- ▶ 3. Once cooperation is established by others, cooperate oneself; stop cooperating if others do not reciprocate; punish noncooperators if feasible.

In addition, one may find at least three other norms

- ▶ 4. Never cooperate.
- ▶ 5. Mimic (1) or (2), but stop cooperating if one can successfully free ride on others.
- ▶ 6. Always cooperate (an extremely rare norm in all cultures).

When many individuals use reciprocity, there is an incentive to acquire a reputation for keeping promises and performing actions with short-term costs but long term net benefits.

Thus, trustworthy individuals who trust others with a reputation for being trustworthy (and try to avoid those who have a reputation for being untrustworthy) can engage in mutually productive social exchanges, even though they are dilemmas, so long as they can limit their interactions primarily to those with a reputation for keeping promises.

A reputation for being trustworthy, or for using retribution against those who do not keep their agreements or keep up their fair share, becomes a valuable asset. In an evolutionary context, it increases fitness in an environment in which others use reciprocity norms. Similarly, developing **trust** in an environment in which others are trustworthy is also an asset

Let us start with a scenario that should be conducive to cooperation - a small group of ten farmers who own farms of approximately the same size. These farmers share the use of a creek for irrigation that runs by their relatively flat properties. They face the problem each year of organizing one collective workday to clear out the fallen trees and brush from the prior winter. All ten expect to continue farming into the indefinite future.

Let us assume that the creek delivers a better water supply directly in response to how many days of work are completed. All farmers have productive opportunities for their labor that return more at the margin than the return they would receive from their own input into this effort. Thus, free riding and hoping that the others contribute labor is **objectively attractive**.

The value to each farmer, however, of participation in a successful collective effort to clear the creek is greater than the costs of participating.

Given the small size of the group, its symmetry, and the relatively low cost of providing the public good, combined with the relatively long time horizon, we can predict with some confidence that a large proportion of individuals facing such a situation will find a way to cooperate and overcome the dilemma. Not only does the evidence from experimental research support that prediction, but also substantial evidence from the field is consistent with this explanation (see B. Ostrom n.d.).

Now let us examine how some structural variables affect the likelihood of collective action (see Figure 3). As a small group, it would be easy for them to engage in face-to-face communication. Since their interests and resources are relatively symmetric, arriving at a fair, contingent agreement regarding how to share the work should not be too difficult.

One simple agreement that is easy to monitor is that they all work on the same day, but each is responsible for clearing the part of the creek going through his or her property. Conformance to such an agreement would be easy to verify. While engaged in discussions, they can reinforce the importance of everyone participating in the workday. In face-to-face meetings, they can also gossip about any one who failed to participate in the past, urge them to change their ways, and threaten to stop all labor contributions if they do not "shape up."

Laboratory experiments provide evidence that a substantial proportion of individuals use reciprocity norms even in the very short-term environments of an experiment (McCabe, Rassenti, and Smith 1996).

Some evidence comes from experiments on ultimatum games. In such games, two players are asked to divide a fixed sum of money. The first player suggests a division to the second, who then decides to accept or reject the offer. If the offer is accepted, then the funds are divided as proposed. If it is rejected, then both players receive zero.

The predicted equilibrium is that the first player will offer a minimal unit to the second player, who will then **accept anything more than zero**. This prediction has repeatedly been **falsified**.

Subjects assigned to the first position tend to offer substantially more than the minimum unit. They frequently offer the "fair" division of splitting the sum. Second movers tend to reject offers that are quite small. The acceptance level for offers tends to cluster around different values in diverse cultures (Roth et al. 1991).

Given that the refusal to accept the funds offered contradicts a basic tenet in the complete model of rationality, these findings have represented a major challenge to the model's empirical validity in this setting.

Several hypotheses have been offered to explain these findings, including a "punishment hypothesis" and a "learning hypothesis." The punishment hypothesis is in essence a reciprocity argument. In contrast to adaptive learning, punishment attributes a motive to the second mover's rejection of an unequal division asserting that it is done to punish the first mover for unfair treatment.

This propensity toward negative reciprocity is the linchpin of the argument. Given this propensity first movers should tend to shy away from the perfect equilibrium offer out of fear of winding up with nothing (Abbink et al. 1996). Abbink and his colleagues designed an experiment in which the prediction of the learning and punishment hypotheses is clearly different and found strong support for the punishment hypothesis.

"We found that second movers were three times more likely to reject the unequal split when doing so punished the first mover than when doing so rewarded the first mover" (Abbink et al. 1996, 15-6). Consequently, second movers do appear to punish first movers who propose unfair divisions.

What the research on social dilemmas demonstrates is a world of **possibility** rather than of **necessity**. We are neither trapped in inexorable tragedies nor free of moral responsibility for creating and sustaining incentives that facilitate our own achievement of mutually productive outcomes. We cannot adopt the smug presumption of those earlier group theorists who thought groups would always form whenever a joint benefit would be obtained. We can expect many groups to fail to achieve mutually productive benefits due to their lack of trust in one another or to the lack of arenas for low-cost communication, institutional innovation, and the creation of monitoring and sanctioning rules (V. Ostrom 1997).

V. Ostrom is probably her husband

Nor can we simply rest assured that only one type of institution exists for all social dilemmas, such as a competitive market, in which individuals pursuing their own preferences are led to produce mutually productive outcomes. While new institutions often facilitate collective action, the key problems are to design new rules, motivate participants to conform to rules once they are devised, and find and appropriately punish those who cheat. Without individuals viewing rules as appropriate mechanisms to enhance reciprocal relationships, no police force and court system on earth can monitor and enforce all the needed rules on its own. Nor would most of us want to live in a society in which police were really the thin blue line enforcing all rules.

Human history teaches us that autocratic governments often wage war on their own citizens as well as on those of other jurisdictions. Democracies are characterized by the processing of conflict among individuals and groups without resort to massive killings. Democracies are, however, themselves fragile institutions that are vulnerable to manipulation if citizens and officials are not vigilant (V. Ostrom 1997). For those who wish the twenty-first century to be one of peace, we need to translate our research findings on collective action into materials written for high school and undergraduate students. All too many of our textbooks focus exclusively on leaders and, worse, only national - level leaders.

Students completing an introductory course on American government, or political science more generally, will not learn that they play an essential role in sustaining democracy. Citizen participation is presented as contacting leaders, organizing interest groups and parties, and voting. That citizens need additional skills and knowledge to resolve the social dilemmas they face is left unaddressed. Their moral decisions are not discussed.

We are producing generations of cynical citizens with little trust in one another, much less in their governments. Given the **central role of trust** in solving social dilemmas, we may be creating the very conditions that undermine our own democratic ways of life. It is ordinary persons and citizens who craft and sustain the workability of the institutions of everyday life. We owe an obligation to the next generation to carry forward the best of our knowledge about how individuals solve the multiplicity of social dilemmas large and small-that they face.

Thank you for listening!