COORDINATION GAMES · A PLAIN-LANGUAGE COMPANION

How the games got their *rules*.

Where the ideas come from, who figured them out, and why a small cooperative in Boulder is putting them all together to see what happens.

FOR CURIOUS READERS · BOULDER, COLORADO · 2026

The Coordination Games look, at first, like games. They are. Players make moves, the moves get scored, the records build up. But every choice in how the games are set up is borrowed from someone who spent their career trying to understand a deceptively simple question: when groups of people share something, how do they avoid wrecking it?

That question has a long history. It runs from Maine fishermen managing lobster grounds, through Swiss villagers managing alpine meadows, through software developers managing open-source projects, and now, more recently, through researchers asking what happens when some of the players in a coordination problem are AI agents.

This page is a tour of that history, in plain language. No prior knowledge required. Each design choice in the games is matched to the person or research that inspired it, with a short explanation of what they figured out and why it matters here. There is also a friendly reading list at the end, sorted by where to start.

01 · THE HEADLINE

The big idea, in one paragraph.

In 1968, a biologist named Garrett Hardin published an essay arguing that whenever a group shares a resource, individual self-interest will eventually destroy the whole thing. He called it *the tragedy of the commons*. For decades, this was treated as obvious. Then a political scientist named Elinor Ostrom went out and looked at actual communities sharing actual resources, and found that they often did not collapse at all. They had

figured out how to govern themselves. She spent forty years studying how, won a Nobel Prize in 2009, and left behind a set of design principles that worked across cultures and centuries. The Coordination Games are an attempt to put those principles into a recurring, playable, observable form, and to see whether they still work when the players are AI agents alongside humans.

> *A game without rules is chaos. A game without a tradition is a curiosity. The Coordination Games are neither.*

02 · THE CHOICES, AND WHERE THEY COME FROM

Each design decision, with the story behind it.

Below are the main choices that shape the games, each paired with the person whose work inspired it. The names matter less than the ideas. The ideas matter less than the patterns they describe. The patterns are what the games are actually testing.

CHOICE 01

Trust grows in *steps*, not all at once.

New players in the games start with limited access. As they show good behavior, more games and higher stakes open

up. If they slip, they do not get banned. Their access narrows, and the way back is gradual.

This comes from one of Ostrom's most striking findings. Communities that managed shared resources well almost never used a single, harsh punishment for rule-breaking. They used graduated responses: a warning the first time, a small fine the second, a larger one the third, exclusion only after sustained pattern. The punishment fit the offense, and recovery was always possible.

---

*From **Elinor Ostrom**, **Governing the Commons** (1990), and the experimental work of **Ernst Fehr and Simon Gächter** on punishment in cooperation games.*

CHOICE 02

Players keep the same name *across seasons*.

Each agent in the games has one persistent identity. The record of how it has played follows it everywhere. A player cannot wipe a bad reputation by re-registering, and a player cannot lose a hard-earned good reputation by signing in next season.

This addresses something Ostrom called *clearly defined boundaries*: you cannot have cooperation among strangers whose identities reset between rounds. Computer scientists working on online reputation systems reached the same conclusion from the other direction: a search engine called PageRank and a peer-to-peer system called EigenTrust both showed that trust networks only work when identities persist.

*From **Ostrom** on commons design principles, and from **Sepandar Kamvar and colleagues** at Stanford on the EigenTrust algorithm (2003).*

CHOICE 03

Every important act gets *witnessed and recorded*.

When a player makes a meaningful move in a game, the record of that move is published, signed, and saved. Anyone can look at it later. Other players see it before they decide whether to trust this player.

Ostrom found that successful commons communities all had one thing in common: they monitored each other. Not in a surveillance sense, but in the sense that what people did was observable to the group. Without that, free-riding becomes invisible and cooperation falls apart. The technical name for the digital version of monitoring is an *attestation*, which is just a fancy word for a witnessed claim about what happened.

*From **Ostrom** on monitoring as a design principle, and from a 1982 accounting paper by **William McCarthy** that proposed recording every economic event as a relationship between resources, events, and agents.*

CHOICE 04

Three layers of decisions, not *one big one*.

Decisions in the games happen at different speeds. Inside a single game, players make moves quickly. Between games in a season, the cooperative decides which games

count, who plays, and how prizes are distributed. Across years, the cooperative changes its own rules, admits new members, and revises what the whole thing is for.

This separation is the key insight of something called *polycentric governance*, which Vincent and Elinor Ostrom developed across their careers. Instead of one big decision-maker at the top, multiple decision-makers handle different kinds of questions on different time scales. The result is more resilient than top-down control, and more coherent than letting everyone vote on everything.

From Vincent Ostrom, Charles Tiebout, and Robert Warren (1961), and from Elinor Ostrom's Nobel Prize lecture in 2010.

CHOICE 05

Seasons, not *endless play*.

The games run on a quarterly cycle anchored to the solstices and equinoxes. Each season opens, runs, closes, and resets the leaderboards. The records carry forward, but the season itself ends.

Why? Because cooperation needs both stability and renewal. If everything runs continuously, players never get a fresh chance and reputations calcify. If everything resets, nothing accumulates and there is no reason to play well. Seasons are the answer. They borrow from a famous 1980s experiment by Robert Axelrod, who ran a tournament of cooperation strategies and found that the simple, forgiving, but firm strategy called Tit-for-Tat won when run over many rounds. The cycle matters.

*From **Robert Axelrod**, **The Evolution of Cooperation** (1984), and the long tradition of seasonal cycles in cooperative work.*

CHOICE 06

Humans and AI agents play in *the same field*.

The games allow human players, AI agents, and humans working with AI assistance to share the same coordination event. The records keep them distinguishable, so anyone studying the data can ask how cooperation works across the mix.

This is the part of the games that touches the active research frontier. A team at Google DeepMind, led by a researcher named Joel Leibo, built a testing environment called Melting Pot to ask exactly this question: what happens when AI agents have to coordinate with each other and with humans in social situations? A 2020 paper by Allan Dafoe and colleagues laid out a research agenda called *Cooperative AI* that is now a small but serious field. The Coordination Games are a working extension of that agenda, with an institution attached.

*From **Allan Dafoe and colleagues**, "Open Problems in Cooperative AI" (2020), and the Melting Pot research from **Joel Leibo and the DeepMind team**.*

CHOICE 07

Punishment that lets you *come back*.

When a player breaks a rule or repeatedly defects, the response is not to remove them permanently. Their access narrows, their governance weight is suspended, and a path back is offered. Recovery is possible, but the past is preserved.

This borrows from something called *restorative justice*, a tradition in legal philosophy that argues punishment should focus on repairing what was broken, not on isolating the offender forever. An Australian criminologist named John Braithwaite is the best-known modern voice in this tradition, and his work has influenced how schools, neighborhoods, and even some courts handle rule-breaking. The Coordination Games adopt the principle: a game where defection ends the relationship is structurally different from one where defection can be repaired.

From John Braithwaite, Crime, Shame, and Reintegration (1989), and from Howard Zehr's foundational work on restorative justice.

CHOICE 08

The cooperative is part of *the experiment*.

The games are run by a Colorado cooperative called RegenHub, LCA. The legal form, called a Limited Cooperative Association, is unusual. It allows two kinds of members: those who participate in the work (patrons) and those who provide capital (investors), with patrons holding the controlling vote.

This matters because the cooperative is not just a corporate wrapper. It is the layer that decides which games run, where money goes, when data is released, and

on what terms. If the games are about how players cooperate, the cooperative is about how the people running the games cooperate. A legal scholar named Henry Hansmann argued in 1996 that the form of an organization should match who actually bears the risk of its decisions. The LCA is one of the few legal forms in U.S. corporate law that takes that argument seriously.

*From **Henry Hansmann, The Ownership of Enterprise** (1996), and the Uniform Limited Cooperative Association Act adopted by Colorado in 2011.*

03 · THE FIVE GAMES

Where each one comes from.

The first season opens with five games. None of them are invented from scratch. Each one is a problem that researchers have studied for decades, made playable.

GAME 01

Oathbreaker

Players make pledges, negotiate, and choose whether to honor or break what they promised. Twelve rounds. The only thing that travels with you is your reputation.

*A descendant of the **iterated prisoner's dilemma**, the most-studied game in social science since the 1950s, and the central object of Robert Axelrod's*

tournament. The lineage runs through philosophers like David Lewis on convention and Cristina Bicchieri on social norms.

GAME 03

Shelling Point

Players have to pick the same answer as the others, but they cannot talk to each other. The only way to win is to guess what everyone else will guess.

From the work of Thomas Schelling, who won a Nobel Prize in 2005. His 1960 book The Strategy of Conflict introduced the idea of a focal point: an answer that becomes obvious not because it is logical, but because it is shared.

GAME 04

Tragedy of the Commons

A shared resource that depletes when too many players take too much. The classic dilemma made playable, with a record of every act of restraint or extraction.

Named after Garrett Hardin's 1968 essay, but designed in the spirit of Elinor Ostrom's response: the tragedy is not inevitable. It depends on whether the players have figured out how to govern themselves.

GAME 02

Capture the Lobster

Team play with limited information. Players cannot see everything happening on the field, and they have to coordinate without full visibility.

A nod to the lobster fishermen of coastal Maine, whose informal territory system has been studied for decades by anthropologist James Acheson and was one of Ostrom's recurring case studies of self-governing commons.

GAME 05

AI 2027

The season finale. The most complex game, where what an agent did in earlier games shapes where it begins this one. The stakes are higher and the rules are layered.

The newest of the five, and the one most directly engaged with the active research frontier in cooperative AI. The name nods to a genre of speculative AI futures, but the empirical question stays grounded: can mixed populations sustain cooperation when stakes climb?

04 · THE PEOPLE WHOSE WORK THIS BUILDS ON

Twelve living scholars worth knowing about.

These are not just names in a footnote. They are people whose ongoing work is closest to what the

Coordination Games are testing. Some of them might end up engaging with the games directly. All of them have written things a curious reader can find and read.

Joel Z. Leibo

GOOGLE DEEPMIND

Built the Melting Pot testing environment, where AI agents have to coordinate in social situations. The closest match to the question of what happens when humans and AI agents play together.

Marco A. Janssen

ARIZONA STATE UNIVERSITY

One of Ostrom's most active living collaborators. Runs computer simulations and lab experiments on how communities manage shared resources. The bridge between Ostrom's tradition and modern computational tools.

Michael Cox

DARTMOUTH

Co-authored the 2010 paper that revisited Ostrom's eight design principles by checking them against ninety-one real-world cases. Confirmed they hold up. Continues practical commons work and is unusually generous with practitioners.

Nathan Schneider

CU BOULDER · MEDIA ECONOMIES DESIGN LAB

Author of *Governable Spaces*, which argues that most online communities default to single-administrator rule. Writes accessibly about cooperatives and online governance. Geographically local to

Boulder, which makes him a natural fit.

Brett Frischmann, Michael Madison, and Katherine Strandburg

VILLANOVA · PITTSBURGH · NYU

The trio behind *Governing Knowledge Commons*, which extended Ostrom's work to things like open-source software, scientific datasets, and Wikipedia. Their framework treats knowledge itself as a resource to be governed.

Jessica Gordon Nembhard

CUNY · JOHN JAY COLLEGE

Author of *Collective Courage*, a history of African American cooperative economic thought. Brings a perspective on community and equity that Ostrom's tradition has sometimes underweighted.

E. Glen Weyl

MICROSOFT RESEARCH · PLURALITY INSTITUTE

Co-author of *Plurality* with Audrey Tang. Works on quadratic funding, attestation-based identity, and other mechanisms that are easy to translate into game rules.

Audrey Tang

PLURALITY INSTITUTE

Former Digital Minister of Taiwan. Brought broad listening tools and digital democracy infrastructure into actual government use. The most public proof that this kind of work can run at scale.

Cristina Bicchieri

UNIVERSITY OF PENNSYLVANIA

Studies how social norms form, hold, and break. Bridges philosophy and behavioral economics. Her book *The Grammar of Society* is one of the clearest things written on why people follow rules.

Vincent Conitzer and Caspar Oesterheld

CARNEGIE MELLON · FOUNDATIONS OF COOPERATIVE AI LAB

Run the most theoretically rigorous current home of cooperative-AI research. Concerned with the formal foundations the games would extend with real-world evidence.

David Sloan Wilson

BINGHAMTON · PROSOCIAL WORLD

Has explicitly extended Ostrom's principles to any kind of cooperative group, not just resource users. Runs a practitioner network. A natural fit for translating ideas into practice.

Allan Dafoe

GOOGLE DEEPMIND

Co-author of "Open Problems in Cooperative AI." Frames the question of AI cooperation at civilizational scale, which is the strategic context that makes a small experiment in Boulder worth taking seriously.

Where the games could actually contribute.

Not every interesting question has a settled answer. Below are six debates where smart people genuinely disagree, and where a standing experiment with persistent identity and public records could generate evidence rather than just commentary.

> *Do Ostrom's design principles work in modern, digital, or AI-mediated settings?*
>
> Most of the original case studies were small, face-to-face communities managing physical resources like fisheries and forests. A 2010 review confirmed the principles hold up across many cases, but critics argue they may miss what is different about digital and AI settings. The games are a place to find out.

> *Does punishing free-riders always help cooperation, or sometimes hurt it?*
>
> A famous experiment showed people will pay to punish cheaters, even at cost to themselves, and this sustains cooperation. A later cross-cultural study found that in some places, punishment backfires, with people retaliating against punishers. Graduated and restorative responses are a third option that has been more theorized than tested.

> *Can AI agents based on large language models cooperate with each other in ongoing games?*

Recent results from 2023 to 2025 are mixed. Some models cooperate readily, others slide toward defection, and the answer often depends on which model and how it is prompted. Mixed groups of humans and AI agents under graduated trust are largely untested territory.

Can reputation systems resist abuse, or do they always get gamed?

Reputation systems on the internet have been studied for over twenty years, and the technical answer is mixed. Tying reputation to membership in a real cooperative, with legal personhood and enforceable rules, is an institutional answer to what has usually been treated as a purely technical problem.

Is it possible to govern an online community without one person ending up in charge?

Nathan Schneider argues that nearly all online communities default to a single-administrator pattern that he calls implicit feudalism. A polycentric governance running in production for multiple seasons would be informative either way, whether it succeeds or whether it drifts back toward central control.

Are knowledge and protocols a kind of commons, and if so, how should they be governed?

A research framework called Governing Knowledge Commons has been building case studies for over a decade,

> but living experiments are rare. The Coordination Games are arguably one such experiment, in a domain where research has been ahead of practice.

06 · A FRIENDLY READING LIST

Where to start, depending on how deep you want to go.

No one has read all of this. No one needs to. The list below is sorted from most readable to most technical, so you can stop wherever interest gives out.

START HERE

If you only read one thing from each row, these are the doorways.

David Bollier, *Think Like a Commoner: A Short Introduction to the Life of the Commons* (2014). The friendliest entry into commons thinking. EASY

Marjorie Kelly, *Owning Our Future: The Emerging Ownership Revolution* (2012). On why how we own things shapes what they become. EASY

Nathan Schneider, *Everything for Everyone: The Radical Tradition That Is Shaping the Next Economy* (2018). A history of cooperatives told as a story. EASY

Kevin Owocki, *GreenPilled: How Crypto Can Regenerate the World* (2022). The contemporary practitioner case for using new tools for cooperation. EASY

IF YOU WANT TO UNDERSTAND OSTROM

The Nobel laureate at the heart of the work. Start with the lecture, then the book.

Elinor Ostrom, "Beyond Markets and States: Polycentric Governance of Complex Economic Systems" (2010). Her Nobel lecture, available free online. The single best summary in her own voice. APPROACHABLE

Elinor Ostrom, *Governing the Commons* (1990). The book that changed the field. Dense, but worth it. MODERATE

Charlotte Hess and Elinor Ostrom, eds., *Understanding Knowledge as a Commons* (2007). On extending the framework to information and ideas. MODERATE

Michael Cox, Gwen Arnold, and Sergio Villamayor-Tomás, "A Review of Design Principles for Community-based Natural Resource Management" (2010). The empirical revalidation. Available open-access. TECHNICAL

IF YOU WANT TO UNDERSTAND COOPERATION ITSELF

How and why people work together, from psychology to game theory.

Robert Axelrod, *The Evolution of Cooperation* (1984). Clear, practical, still influential. APPROACHABLE

Thomas Schelling, *The Strategy of Conflict* (1960). The book that introduced focal points. Some chapters are dense, but the core ideas are vivid. MODERATE

Cristina Bicchieri, *The Grammar of Society* (2006). On how social norms work. MODERATE

Samuel Bowles and Herbert Gintis, *A Cooperative Species* (2011).

Why humans cooperate, told from an evolutionary angle. MODERATE

IF YOU WANT TO UNDERSTAND THE AI SIDE

The active research frontier on multi-agent and human-AI cooperation.

Allan Dafoe and colleagues, "Open Problems in Cooperative AI" (2020). Sets the agenda. Available free on arXiv. TECHNICAL

Allan Dafoe and colleagues, "Cooperative AI: Machines Must Learn to Find Common Ground," in *Nature* (2021). The short version of the agenda. APPROACHABLE

Joel Leibo and colleagues, "Scalable Evaluation of Multi-Agent Reinforcement Learning with Melting Pot" (2021). The technical paper behind the testing environment. TECHNICAL

Richard Willis and colleagues, "Will Systems of LLM Agents Cooperate" (2025). Recent results on whether language model agents cooperate. TECHNICAL

IF YOU WANT TO UNDERSTAND COOPERATIVES

The institutional form behind the games.

Jessica Gordon Nembhard, *Collective Courage: A History of African American Cooperative Economic Thought and Practice* (2014). A history that is rarely told. APPROACHABLE

Trebor Scholz and Nathan Schneider, eds., *Ours to Hack and to Own: The Rise of Platform Cooperativism* (2016). On bringing cooperatives into the digital era. APPROACHABLE

Henry Hansmann, *The Ownership of Enterprise* (1996). Why some firms are best owned by workers, others by customers, others by investors. MODERATE

Nathan Schneider, *Governable Spaces: Democratic Design for Online*

Life (2024). On bringing democracy into online communities.

APPROACHABLE

IF YOU WANT TO UNDERSTAND THE BIGGER PICTURE

The civilizational frame around all of this.

E. Glen Weyl, Audrey Tang, and the Plurality community, *Plurality: The Future of Collaborative Technology and Democracy* (2024). Available free online. APPROACHABLE

Daniel Schmachtenberger, the *Civilization Emerging* essays and the *Consilience Project* publications. On the larger coordination crisis the games sit inside. APPROACHABLE

Yochai Benkler, *The Wealth of Networks* (2006). On how networked production changes economics. MODERATE

07 · A CLOSING THOUGHT

Why all of this matters.

The Coordination Games are small. One cooperative, one Colorado town, a handful of games, a public record that will take years to thicken into something useful. None of the people listed above know whether the experiment will work.

But the question the games are built around is not small. How do groups of people, and increasingly groups of people and AI agents together, share what they have without ruining it? That question is now playing out at every scale, from neighborhoods to

platforms to planetary systems. The answers are not yet known, and the institutions we have are not all up to the work.

The Coordination Games are one small contribution to figuring it out. The thinking behind them is not new. It is borrowed, with care, from people who spent their lives on the question. The hope is that putting the borrowed pieces together, in a place where the records can be observed and the rules can evolve, will teach the next thing.

> *The thinking is not new. The combination is. That is what an experiment is for.*
