Putting Behavior into International Cooperation

Dirk Messner  
German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE), Tulpenfeld 6, 53113 Bonn, Germany  
Dirk.Messner@die-gdi.de

Alejandro Guarin  
German Development Institute / Deutsches Institut für Entwicklungspolitik (DIE), Tulpenfeld 6, 53113 Bonn, Germany  
Alejandro.Guarin@die-gdi.de

Daniel Haun  
Department of Early Child Development and Culture, Leipzig University, Marschnerstr. 29, 04109 Leipzig, Germany  
daniel.haun@uni-leipzig.de

Abstract

Cooperation among countries to deal with climate change and other global challenges seems almost impossible. However, recent research in the behavioral and evolutionary sciences suggests that people cooperate willingly and perhaps even naturally. Why the contradiction? In this paper we call for the need to think about international relations beyond the assumptions of rational choice and narrow self-interest. We suggest that seemingly intractable international challenges are not only the result of power struggles, but may also reflect the under-provisioning of the key enablers of cooperation—what we call the “cooperation hexagon”: reciprocity, trust, we-identity, enforcement, communication, reputation and fairness.

Keywords

1 Introduction

The latest IPCC report leaves very little room for doubt: the effects of human-driven climate change are already being felt, and the consequences will be devastating unless swift and meaningful action is taken to mitigate it.1 Despite the recent milestone agreement reached in Paris in 2015, the commitments of most countries still fall short of what scientists think is needed to avoid irreversible and potentially catastrophic consequences of climate change.2 On the face of it, this is a typical case of a cooperation dilemma: even though all countries would benefit from avoiding catastrophic climate change, few are willing to put up with the costs of preventing it. This lack of cooperation is the obvious result if we assume that people are driven only by their self-interest. But what if this assumption is wrong?

Opening his Nobel Prize Lecture in 2001, George Akerlof called for “the development of a behavioral macroeconomics in the original spirit of Keynes’ General Theory”, which would incorporate assumptions about the “role of psychological and sociological factors, such as cognitive bias, reciprocity, fairness, herding, and social status”.3 Quite simply, Akerlof was advocating for bringing microeconomic behavior into macroeconomics. While microeconomics had benefitted from rich insights about the nature of human psychology and behavior, these insights had not been incorporated into—let alone accepted by—mainstream—macroeconomic theory and models. The work of Akerlof and others has shown that considering important elements of human behavior—what Keynes called “animal spirits”—allows macroeconomics to see the world differently.4 By applying a behavioral lens on macroeconomics not only do previously unexplainable problems come into focus, but a whole set of new problems becomes visible.

Such a formulation of a richer macro theory that incorporates insights of micro behavior is lacking in the field of international relations and in global governance research. Most scholars assume that, in the absence of

---

a global government, nations inevitably interact with each other based mainly or even solely on their self-interest by means of power games.\textsuperscript{5} From this perspective, international cooperation to successfully address critical issues such as climate change, or the development of global governance regimes to manage other global commons is, by definition, very unlikely. This view ultimately is based on assumptions about human behavior which have a long intellectual tradition. On the one hand is a fundamentally pessimistic assumption, stemming from Thomas Hobbes’s notion that the natural state of human beings is war “of every man against every man”\textsuperscript{6} and from Herbert Spencer’s interpretation of natural selection as “survival of the fittest”.\textsuperscript{7} On the other hand, in the lineage of Adam Smith, is the more optimistic view in which human self-interest adds up to a general benefit by means of a self-organizing system.\textsuperscript{8}

Both of these assumptions are at odds with our current knowledge about human behavior. Recent research suggests that cooperation is neither impossible nor automatic: cooperation emerges frequently, provided that certain conditions are in place. In this article, we argue that the weight of this new evidence about human behavior is enough to prompt a new perspective on international cooperation and global governance. Our call is for a more inclusive theory, one that acknowledges power struggles, but that at the same time recognizes that there is a wide space for cooperation and investigates the conditions under which it can occur.

2 What We Know About Small-Scale Cooperation

Advances in psychology, evolutionary biology, economics and other social sciences over the last two decades have shown human behavior in a new light. A wealth of evidence from laboratory and field experiments has shown that, in decision-making games, people routinely cooperate—well beyond the predictions of rational choice theory.\textsuperscript{9} Similarly, careful documentation of the

actual management of common resources has demonstrated that Hardin’s “Tragedy of the commons” is neither widespread nor inevitable.

These results are supported by evidence from comparative psychology and evolutionary anthropology. Research carried out with humans of different ages and cultural backgrounds, as well as with other primates, suggests that such outcomes may be the result of a natural propensity to cooperate. The evidence can be summarized into three core insights. First, children cooperate from a very early age (from ca. 18 months), and enjoy doing so, often without seeking a reward, and presumably before being taught to. Second, other primates show some basic abilities to cooperate, which suggests that cooperation is well established in our evolutionary lineage. Moreover, plausible mechanisms for the evolution of cooperative populations at all levels of organization have been amply demonstrated. Third, some key traits of cooperation appear to hold cross-culturally, suggesting that the predisposition to cooperate is independent of culture.

3 The Cooperation Hexagon

If there is a natural bias, why does cooperation take place in some cases but not in other cases? We have found a remarkable convergence from experimental and in situ research from multiple disciplines on the relatively small set of mechanisms that enable cooperation: reciprocity, trust, communication, reputation, fairness, enforcement and we-identity. We call this set the cooperation hexagon (Figure 1). At the centre is reciprocity, a fundamental prerequisite for cooperation to be sustained in time. Despite our pro-social tendencies, we learn early in our cultural development one of the central norms taught in all societies: don't be a sucker.19 Direct reciprocity—the strategy of tit-for-tat—is one of the crucial mechanisms underlying the evolution of stable cooperative populations.20 And yet we reciprocate within a wide circle of people, many of whom are not directly related to us, and whom we will never see again. Language, gossip, reputation and other mechanisms of control make indirect reciprocity—“I scratch your back and someone will scratch mine”—possible.21 Observations by economists and political scientists of commons management in the field suggest that one of the key ingredients of success is the assurance that contributions will be returned.22 In decision-making experiments such as the ultimatum game, generous offers have been interpreted as involving an expectation of repayment; partners, unsurprisingly, tend to reciprocate rather than exploit this behavior.23 Obviously reciprocity is not always positive: in some contexts in can lead to a downward spiral of anti-cooperation in which even formerly cooperative people stop cooperating.24

The other six elements of the diagram create conditions conducive to reciprocity. Trust can be understood as a belief about the probability of reciprocation25 motivated by an aversion to being betrayed.26 Decision-making

---

21 Martin Nowak and Roger Highfield, Super Cooperators: Evolution, Altruism and Human Behaviour or Why We Need Each Other to Succeed (Edinburgh: Canongate, 2011).
22 Poteete, Janssen, and Ostrom, Working Together.
23 Bowles and Gintis, A Cooperative Species: Human Reciprocity and Its Evolution.
24 Nowak, “Five Rules for the Evolution of Cooperation.”
experiments have shown that the benefits of cooperation are enhanced if individuals are willing to take risks because they trust each other. In small groups where interactions occur repeatedly we are able to keep track of other’s actions to determine their trustworthiness. But in a world full of strangers, we have evolved to rely first on external features by, for example, being biased towards trusting those who look similar to us, even if we later refine our judgments based on experience.

Although cooperation can emerge in anonymous interactions, there is considerable evidence that communication—to share expectations, devise joint strategies and make pledges about future behavior—enhances the chances of cooperative outcomes to persist. Field and laboratory experiments involving public goods have found that allowing for communications significantly increases contributions to the public good. It has been suggested that the

---

importance of communication to sustain cooperation goes beyond information exchange. Communication binds people together because it allows them to share stories and create common narratives; as such, it has been a key ingredient of our species’ cultural evolution. Furthermore, communication is essential for creating trust: we feel bound to stick to our pledges, and others and this creates virtuous cycle of increasing trust among cooperating partners.

In large groups, direct interaction with actual or potential cooperation partners is unlikely. This prevents direct assessments of trustworthiness (via physical signals) or the building of trust through communication. In these cases, we rely on reputation to establish others’ past performance to try to guess how they will behave in the future. Evidence from dilemma experiments suggests that cooperation decreases when participants do not have a chance to build a reputation. But when information flows in the form of news or gossip, reputations synchronize across populations so that everybody knows whom to trust and whom not to. Public acts like shaming or honoring can have an important effect in triggering (or obstructing) cooperation, as has been shown in recent public goods experiments.

For cooperation to succeed, it is often not enough that acts are reciprocated; reciprocation should also be perceived as fair. Inequality aversion has been consistently shown to be an important factor in shaping cooperation, not just...
in humans from an early age\textsuperscript{39} but in other social primates.\textsuperscript{40} In economic experiments, participants modulate their behavior or judge the behavior of others with regard to fairness concerns: they want to be seen as fair and expect others to be fair as well.\textsuperscript{41} For example, in ultimatum games people frequently reject offers that are considered too low (and thus unfair), even if that means leaving empty-handed.\textsuperscript{42} In public goods experiments in the laboratory and the field, great differences in endowments or power have been shown to obstruct communication, hinder the development of trust, and result in poor contributions to the public good.\textsuperscript{43}

Cooperative behavior must sometimes be triggered by some means of enforcement, such a punishment or reward. Enforcement is an external means to ensure reciprocity and typically works in interaction with other mechanisms of the cooperation hexagon. It has been well established in economic experiments that people actively punish cheaters—that is, those who harm the principles of reciprocity and fairness—even if it comes as cost,\textsuperscript{44} and that punishment is often essential to sustain collaboration in public goods games.\textsuperscript{45} To be effective, enforcement frequently plays on the value of reputation: both public shaming\textsuperscript{46} and rewarding\textsuperscript{47} have been shown to be effective means of enhancing cooperation in public goods games.

\textsuperscript{40} Katherine A. Cronin and Angel Sánchez, “Social Dynamics and Cooperation: The Case of Non-human Primates and Its Implications for Human Behavior,” Advances in Complex Systems 15, no. supp01 (June 2012): 1250066.
\textsuperscript{46} Jacquet et al., “Shame and Honour Drive Cooperation.”
Finally, the emergence and stabilization of cooperation is far more likely to be successful if cooperators share a common *we-identity*. Like enforcement, a common identity works by interacting with other of the mechanisms described above. We tend to associate more with those who look or speak like us because it increases the likelihood that they will reciprocate. Economic experiments have shown that, in public goods games, people tend to contribute more when they are among others who are physically or intellectually similar, or if they belong to their in-group—even if that group is defined arbitrarily. Evidence from the management of common pool resources also suggests that groups that are ethnically or culturally different have a much harder time to develop the necessary trust for effective cooperation. A *we-identity* is not just externally determined: through language we learn and build joint narratives that reinforce our sense of belonging—to religions, political parties, nation-states or football teams—that expand our common ground. A common identity also streamlines our social interactions; having shared understandings and assumptions allows us to manage complexity by quickly agreeing on basic things such as what is acceptable or expected in a wide range of situations.

The mechanisms outlined above reinforce and complement each other to create opportunities for cooperation to emerge. The literature suggests that no mechanism alone is sufficient to trigger cooperation, but it is also not clear

---


what specific or minimum combination is enough to do so. Starting from a non-cooperative situation, it is unlikely that all the mechanism will fall into place simultaneously and lead to cooperation. It seems more possible that cooperation starts with one or two of them, progressively building up on their reinforcing strengths. And while the presence of all the elements of the hexagon is very likely conducive to stable cooperative outcomes, it is clear from computer simulations that even cooperative populations can evolve towards non-cooperative states.56

4 Scaling Up: From Cooperative Behavior to International Cooperation

The cooperation hexagon provides a potentially powerful new way of understanding cooperation problems at any scale, including at the international level. The evidence suggests that cooperation plays a role in human behavior well beyond what is allowed by models of exclusively self-interested ‘rational fools’.57 As we have shown above, the emergence of cooperation at interpersonal levels is shaped by the interplay of several different mechanisms. If two people fail to cooperate, it would be wrong to attribute this failure simply to a narrow calculus of costs and benefits. Our argument, which we explain below, is that it would be just as wrong to assume that blockades in international cooperation are exclusively the result of power games and narrow self-interest.

A behavioral lens on international cooperation has not yet been theoretically or empirically developed, but we suggest that the search be conducted in three directions: 1) Unveiling how the mechanisms of the cooperation hexagon work at different scales; 2) studying how cooperation works within the groups of people who represent nations in international institutions, regimes, negotiations and networks; and 3) looking at how the basic mechanisms of cooperation appear—or falter—in the current arrangements of international cooperation amid tectonic power shifts in the world order.

First, we need to better understand the scaling properties of the mechanisms of the cooperation hexagon. We know that cooperation has been an important driver of complexity and structure at all levels of biological organization,58 and integral to the ratcheting-up of knowledge by which we build on what other

56 Nowak, “Five Rules for the Evolution of Cooperation.”
58 Nowak and Highfield, Super Cooperators.
generations have learned—the keystone of human cultural evolution. And yet the question remains how and why the basic building blocks of interpersonal cooperation (i.e. the hexagon) come together to build larger and more complex cooperative institutions. Are reciprocity, communication, trust, reputation, fairness, enforcement and we-identity scale free? It is well known that simple rules of interaction can lead to systems of great size and complexity. However, scaling up cooperation is not simply about adding more people to the equation. It is also about the structures, institutions and systems of representation that allow for effective cooperation at different levels of organization across society. It is clear that the basic mechanisms of cooperation are produced and reproduced differently depending on the type of structure and its scale. For instance, the development of trust among people in an informal neighborhood association is probably different to that of parliamentarians in a national stage. We have very little empirical evidence about these differences. And yet such knowledge might be instrumental in designing international institutions that are conducive to the flourishing of the fundamental elements of cooperation.

Second, it is necessary to investigate how cooperation works at the formal and informal institutions and regimes through which countries cooperate with each other—one of the core concerns of regime theory. Most of what we call international cooperation is actually carried out by relatively small groups of people who interact frequently as representatives of their constituencies at international fora like climate negotiations, the Chemical Weapons Convention or the Financial Action Task Force on Money Laundering. These groups are a unique point of confluence of interpersonal and inter-institutional cooperation, in which members interact simultaneously as individuals and as representatives of the interests of others. The second aspect has been emphasized by regime theory, as it seeks to understand how nations might cooperate in the face of (and despite) self-interest. To be sure, elected representatives have incentives to act in ways that help them secure electoral victo-

ries at home, but this is not the only driver of their behavior. Do they know the other members? Do they trust them? Do they have the same information at hand? These are questions that are likely to structure their daily interaction. The problem of cooperation in international governance institutions is thus not about the size or complexity of the group, but rather about different (and potentially conflicting) drivers and motivations of cooperation. An explicitly behavioral lens on international regimes could help us understand how and why those motivations are (not) reconciled.

Finally, we could understand international cooperation challenges in terms of the underprovisioning of the basic mechanisms of cooperation. Successful examples of effective international cooperation, such as the European Union, are the result of long and fraught efforts to build trust, communication, fairness, and we-identity. Such processes can pay off, unthinkably, even when carried out by former foes. But cooperation in newly emerging international configurations such as the G20 (the informal group of the world’s 20 largest economies) is much more uncertain. Undoubtedly power struggles play a big role in determining how these nations interact with each other. But while this may be an important part of the story, we believe that it is not the whole story. In today’s world old and new global powers must learn to cooperate to address complex global problems such as the financial system, but they have had no experience working together and very little time to build trust, communication and the other crucial mechanisms of the cooperation hexagon. Therefore, rather than seeing international cooperation blockades solely as a natural consequence of power games, the seeming intractability of global challenges could plausibly also be attributed to an underprovisioning of the basic mechanisms of cooperation. Many nations have not yet properly developed the right channels and institutions to develop reciprocity, trust, communication or reputation. In international relations quarrels stem from the perception unfair burden-sharing or the absence of credible punishing mechanisms. The most difficult task of all, the creation of a global we-identity, is still hampered by the cognitive frameworks that emerged in the era of the nation state.

---


Investing in these mechanisms is costly but necessary; and it involves a strategy that is different to containment to manage power struggles.

Putting behavior into international cooperation theories would provide a richer and more accurate picture of how the world works. International relations theories are built on implicit assumptions about human behavior in the lineage of Hobbes; what we know from diverse disciplines and multiple research methods is that these assumptions are incomplete. Perhaps the natural stage of humans is not of inevitable war of all against all, but instead of a basic predisposition for cooperation? We are not proposing an overhaul of current theories of international relations, but rather their further development. This would involve taking the insights from cooperation at the interpersonal level seriously, and adopting more accurate assumptions about human behavior. Such a development would also bring international relations and global governance theories closer to the mainstream of other disciplines such as economics. Beyond the contribution to theory, taking the insights from micro behavior seriously and making the right assumptions about how people behave could help design better and more effective mechanisms and institutions to solve today’s pressing global problems.