Cooperative reciprocity is widespread in international relations

Morgan R. Frank[1], Nick Obradovich[1,2], Lijun Sun[1], Sohan Dsouza[1], Pai-Ju Chang[1], Wei Lee Woon[3], Brad L. LeVeck[4], Iyad Rahwan[1,5*]

[1]Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA
[2]Belfer Center for Science and International Affairs, Kennedy School of Government, Harvard University, Cambridge, MA, USA
[3]Electrical Engineering and Computer Science, Masdar Institute of Science and Technology, Abu Dhabi, UAE
[4]Department of Political Science, University of California, Merced, CA, USA
[5]Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA, USA
*To whom correspondence should be addressed; E-mail: irahwan@mit.edu.

Reciprocity has been shown to stabilize cooperation from the level of microbes all the way up to humans interacting in small groups, but does reciprocity also underlie stable cooperation between larger human agglomerations, such as nation states? Famously, evolutionary models show that reciprocity could emerge as a widespread strategy for achieving international cooperation. However, there is only limited empirical support for this hypothesis. Here, we employ a novel empirical approach, which detects causality in non-linear dynamical systems, to detect reciprocity between many country-pairs in the international system. These reciprocating country-pairs exhibit higher levels of stable co-

operation, but are also more likely to reciprocate conflict if provoked. Together, these findings provide the strongest evidence to date that reciprocity is a widespread mechanism for achieving international cooperation.

The international system lacks a single sovereign capable of enforcing cooperative agreements (*1–3*). Therefore, stable inter-country cooperation relies on countries' self-interest for its propagation (*4*). Because reciprocity in repeated games incentivizes cooperation (*5–8*), even in the absence of external enforcers, reciprocity may provide a crucial explanation for bilateral cooperation across a wide variety of international domains, including trade (*9*), the avoidance of war (*10*), and the environment (*11*). In fact, recent experimental studies show that bilateral reciprocity between pairs of actors can sustain cooperative contributions to a shared public good (*12*). Therefore, bilateral reciprocity may even underlie some instances of multilateral cooperation, such as global emissions agreements.

Moreover, evolutionary models (*5, 8, 13, 14*) and laboratory experiments (*6, 7, 15*) both show that simple strategies of reciprocity, such as Tit-for-Tat (TFT), can become widespread in a population—at least when actors place a sufficiently high value on the payoff to future cooperation (*7, 15*). Because states are long-lived actors that typically interact for indefinite periods of time, scholars have hypothesized that TFT-like reciprocity may similarly emerge as a prevalent strategy in international relations (*8, 10, 16*). If so, cooperative reciprocity should be detectable across a large number of country-pairs.

Yet, despite the central role of reciprocity in theories of international cooperation, there is surprisingly little evidence that reciprocity is actually widespread. Empirical studies in international relations have mainly found quantitative and case-based evidence that TFT (*17–19*) as well as other forms cooperative reciprocity (*16, 20, 21*) exist between a small set of countries. Only a few studies have examined the ubiquity of reciprocity in international relations (*22, 23*), and none of these studies investigate whether the dynamics of reciprocity help stabilize coopera-

tion between country pairs. Therefore, these studies do not indicate whether reciprocity actually represents a mechanism for widespread, stable cooperation. Meanwhile, there is some evidence that frequently occurring misperceptions or unintentional actions can destabilize cooperative reciprocity between states (*24, 25*). These findings cast some doubt on whether reciprocity can act as a stable mechanism for widespread international cooperation, even if reciprocity is known to be effective in other social domains. So, the question remains: how widespread is reciprocity in the international system, and does it stabilize cooperation?

Here, we employ a novel empirical approach to conclusively detect reciprocity based on underlying dynamical causation. Our analysis utilizes the Integrated Crisis Early Warning System (ICEWS), which is a recently available global-scale data set containing the interactions of nations based on nearly 30 million news stories from around the world (*26*). Interactions among nations are categorized according to the Conflict and Mediation Event Observations (CAMEO) (*27*) taxonomy, which associates interaction types with real-valued "Goldstein scores" scores ranging between -10 and $+10$ indicating the conflictive or cooperative nature of the interaction. The highest abstraction of the CAMEO taxonomy, called "Quad Classes," includes verbal cooperation, material cooperation, verbal conflict, and material conflict. Aggregating the interactions of a pair of countries on each day from 1/1/1995 through 12/31/2014 and calculating the weighted average of event scores produces a time series representing the temporal dynamics in the treatment of one country towards another (*28–30*) (examples in Fig. 1A).

Specifically, given a collection of interactions, E_{AB}, detailing nation A's level of cooperation with nation B, the average Goldstein score for the collection is given by

$$GS(E_{AB}) = \sum_{e \in C} g_e \cdot p_{E_{AB}}(e), \tag{1}$$

where C is the set of CAMEO event types, $p_{E_{AB}}(e)$ is the proportion of events in E_{AB} of type e, and g_e is the Goldstein score associated with CAMEO event type e. Collections of interactions

can differ according to relative changes in the valence of interaction types and/or changes in the relative abundance of interaction types (more details in S3).

Using these temporal patterns of cooperation and conflict, to what extent does state A's level of cooperation with state B influence B's level of cooperation with A, and vice versa? Convergent cross mapping (CCM) (*31, 32*) is a new method for detecting dynamical causality, or "influence", from time series, and has been used for causal inference from dynamical systems in ecology (*33*), in empirical studies of social media (*34*), and in empirical studies of neuroscience (*35*). CCM uses the closeness of points in one time series to reconstruct a second time series; if the reconstruction is skillful according to Pearson correlation (typically $CCM(A, B) \geq 0.25$ for noisy empirical data (*31*)), then we conclude that the second time series causally influences the first time series (see S5 for calculation, see S6 for details on varying the CCM influence threshold). As an example, Figure 1B&C demonstrate the attitudes and influence from directed Goldstein time series among European Union nations based on interactions from 1995 to the end of 2014.

Applying CCM to the directed Goldstein time series of country A's level of cooperation with country B, and vice versa, identifies pairs of countries exhibiting "CCM reciprocity" (i.e. $CCM(A, B) \geq 0.25$ and $CCM(B, A) \geq 0.25$). That is, if countries A and B have CCM reciprocity, then country A's treatment of country B "CCM causes" country B's treatment of country A and vice versa. Examples of reciprocating country pairs are often nearby spatially, such as Russia and Ukraine which share a border, but the reciprocity between China and the United Kingdom demonstrate how an increasingly connected world allows influence to span distance as well (see Fig. 1D). In total, we detect 47 country pairs exhibiting CCM reciprocity.

Mathematically, CCM reciprocity is not necessarily direct reciprocity since CCM influence may not preserve valence (i.e. does cooperation breed cooperation in kind?). However, compared to non-reciprocating country pairs, CCM reciprocity indicates country pairs that are more

likely to cooperate regardless of recent interactions (see Fig. 2A & B) and are more likely to engage in conflict in response to recent conflict (see Fig. 2D). Furthermore, country pairs in reciprocating relationships are more likely to mirror the specific type of cooperation or conflict that is directed at them (see Fig. 3A-C). This preservation of valence indicates that CCM reciprocity serves as a plausible proxy for direct reciprocity. It is also important that valence is preserved for both verbal interactions (Fig. 3A), as well material cooperation and conflict (Fig. 3B-C). This suggests that our results are not purely explained by "cheap talk", which is abundant in international relations (*36, 37*), but also rely on more "costly" material interactions. Additional comparisons are provided in S6, including response to recent cooperation or conflict compared to aggregate cooperation/conflict on a pair-by-pair basis, and response by Quad Class interaction type.

On aggregate, country pairs with greater shared influence (i.e. $(CCM(A,B)+CCM(B,A))/2$) also exhibit a greater correlation between their directed Goldstein time series (see Fig. 3D). Similar to our results in Fig. 3A-C, this relationship is broadly consistent with models of Tit-for-Tat reciprocity, where players respond in kind to each other's prior actions. However, CCM influence is also not the same thing as a simple correlation in each country's level of cooperation towards the other, as there are country-pairs that exhibit high levels of CCM influence, but have a relatively low correlation between their directed Goldstein time series. One reason this may occur is if countries use more complex strategies to determine whether to defect or return to cooperation. For example, if countries use forgiving strategies that wait for several transgressions before retaliating (*13, 15, 24, 25*), this can weaken the correlation between their directed Goldstein time series.

A willingness to sustain cooperation and forgive transgressions can be crucial for the evolution of cooperation (*13, 15*)—especially when there is a chance that players will misperceive the intent or nature of each other's actions (*15*). Such strategies may therefore be especially

important in international cooperation, where there is a high risk that such misperceptions will occur (*24, 25*). It is therefore interesting to see that reciprocating country-pairs are, in fact, more willing to sustain cooperation (see Fig. 2A) and more likely to return to cooperating after conflictive interactions (see Fig. 2B). These findings demonstrate that reciprocity is characterized by higher levels of forgiveness even when faced with non-cooperation.

Each of the findings above should temper prior skepticism about whether reciprocity represents an widespread mechanism for the emergence of international cooperation. (*38, 39*). In other cases, influence may essentially be one-way, Unlike the actions of individuals, international relations require the coordination of strategies among all actors in a given country (e.g. businesses, government agencies, etc.). Despite this, many results about individual cooperation from the field of cognitive science appear in international relations as well. Therefore, analogous methods for international cooperation can be employed to facilitate new and long lasting cooperative international institutions, and analogous understanding and diagnoses from cooperation in smaller-scale systems can be applied to breakdowns in cooperation in the international arena.

References and Notes

1. K. A. Oye, *World Politics* **38**, 1 (1985).

2. C. Lipson, *World Politics* **37**, 1 (1984).

3. T. Hobbes, *Leviathan* (A&C Black, 2006).

4. J. D. Fearon, *International organization* **52**, 269 (1998).

5. A. J. Stewart, J. B. Plotkin, *Proceedings of the National Academy of Sciences* **110**, 15348 (2013).

6. P. D. Bó, *The American Economic Review* **95**, 1591 (2005).

7. P. D. Bó, G. R. Fréchette, *The American Economic Review* **101**, 411 (2011).

8. W. D. Hamilton, R. Axelrod, *Science* **211**, 1390 (1981).

9. C. Rhodes, *International Organization* **43**, 273 (1989).

10. R. M. Axelrod, *The evolution of cooperation* (Basic books, 2006).

11. M. M. Bechtel, K. F. Scheve, *Proceedings of the National Academy of Sciences* **110**, 13763 (2013).

12. O. P. Hauser, A. Hendriks, D. G. Rand, M. A. Nowak, *Scientific Reports* **6**, 36079 (2016).

13. M. A. Nowak, K. Sigmund, *Nature* **355**, 250 (1992).

14. R. Axelrod, *et al.*, *The Dynamics of Norms* pp. 1–16 (1987).

15. D. Fudenberg, D. G. Rand, A. Dreber, *The American Economic Review* **102**, 720 (2012).

16. R. O. Keohane, *International Organization* **40**, 1 (1986).

17. H. Milner, *World Politics* **44**, 466 (1992).

18. P. Ekman, A. I. Hadjikhani, A. Pajuvirta, P. Thilenius, *International Business Review* **23**, 1049 (2014).

19. D. Tingley, M. Tomz, *Comparative Political Studies* p. 0010414013509571 (2013).

20. J. S. Goldstein, *International Studies Quarterly* **35**, 195 (1991).

21. J. S. Goldstein, J. C. Pevehouse, D. J. Gerner, S. Telhami, *Journal of Conflict Resolution* **45**, 594 (2001).

22. S. Minhas, P. D. Hoff, M. D. Ward, *Journal of Peace Research* **53**, 491 (2016).

23. P. D. Hoff, *The annals of applied statistics* **9**, 1169 (2015).

24. R. Jervis, *Perception and misperception in international politics* (Princeton University Press, 2015).

25. J. Wu, R. Axelrod, *The Journal of Conflict Resolution* **39**, 183 (1995).

26. E. Boschee, *et al.*, ICEWS coded event data (2016).

27. J. S. Goldstein, *Journal of Conflict Resolution* **36**, 369 (1992).

28. B. O. Fordham, *Journal of Politics* **67**, 132 (2005).

29. S. M. Shellman, *Political Analysis* **12**, 97 (2004).

30. J. E. Yonamine, *Penn State University: Working Paper* (2011).

31. G. Sugihara, *et al.*, *Science* **338**, 496 (2012).

32. H. Ye, E. R. Deyle, L. J. Gilarranz, G. Sugihara, *Scientific Reports* **5** (2015).

33. A. T. Clark, *et al.*, *Ecology* **96**, 1174 (2015).

34. C. Luo, X. Zheng, D. Zeng, *Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint* (IEEE, 2014), pp. 260–263.

35. L. Heskamp, *et al.*, *Cerebrovascular Diseases* **35**, 15 (2013).

36. J. D. Fearon, *American Political Science Review* **88**, 577 (1994).

37. R. Jervis, *The logic of images in international relations* (Columbia University Press, 1989).

38. M. Tomz, *Reputation and international cooperation: sovereign debt across three centuries* (Princeton University Press, 2012).

39. G. W. Downs, D. M. Rocke, P. N. Barsoom, *International Organization* **50**, 379 (1996).

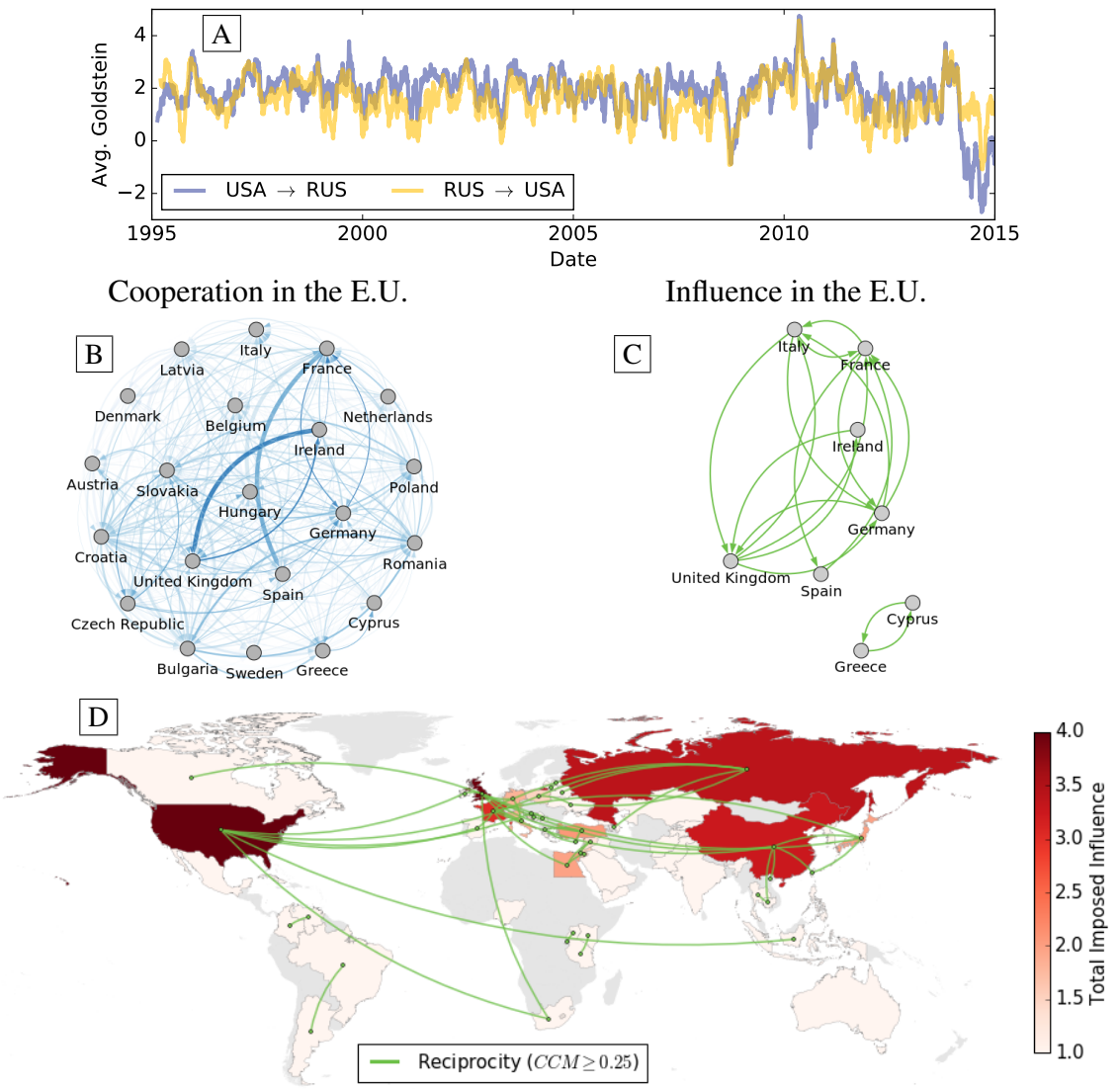

Figure 1: Mapping cooperation, influence, and reciprocity in the European Union (EU) and around the world. **(A)** The directed Goldstein time series for the United States-Russia relationship. The time series have been smoothed using a 30 day moving average for visualization purposes. **(B)** A network representing overall attitude of EU countries, the amount of interaction between nations (darker edges indicate more actions), and how cooperative interactions are on average (edge width). **(C)** A network representation of influence amongst EU nations (i.e. $CCM(A, B) \geq 0.25$). **(D)** Countries are colored according to their total imposed influence on others and green lines connect pairs of countries exhibiting reciprocity (see S5.1). Grey countries had insufficient data for CCM analysis.

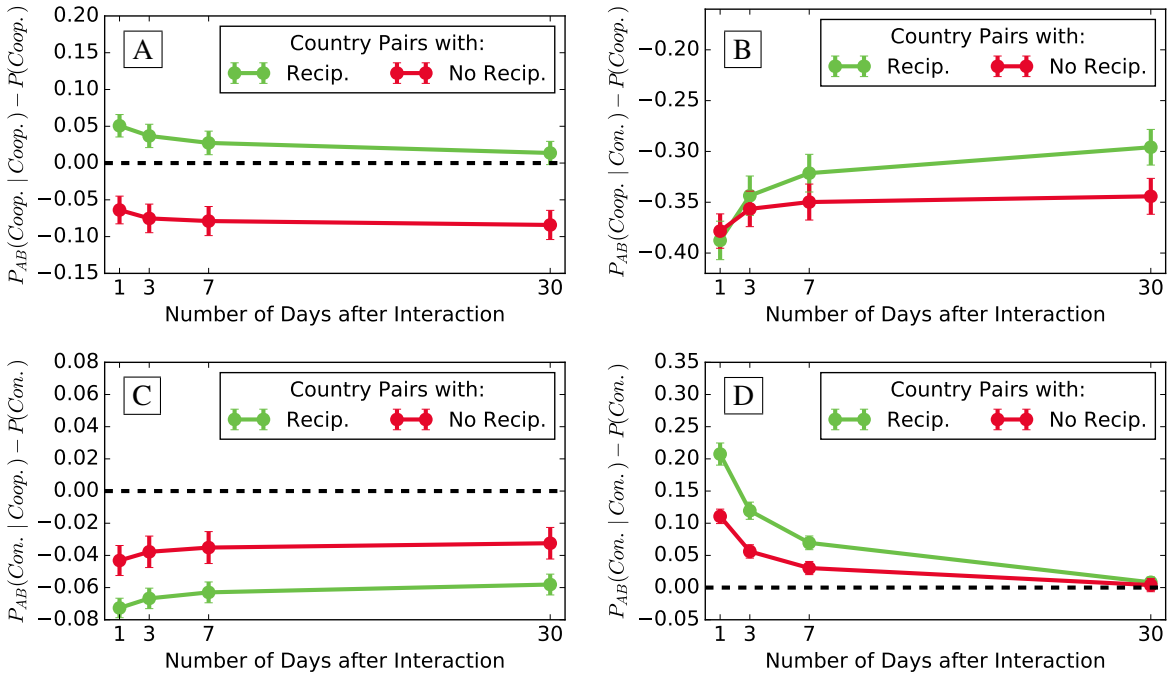

Figure 2: Country pairs exhibiting reciprocity are cooperative on average, but reciprocate conflict. Given an observation of cooperation (left) or conflict (right), reciprocating country-pairs are more likely to cooperate ((**A) & (B)**) regardless of recent interaction, less likely to conflict given recent cooperation **(C)**, but more likely to reciprocate conflict **(D)** in the cumulative interactions of the following day, three days, and seven days (x-axis). Each point represents the average rate of cooperation or conflict between countries A and B, denoted P_{AB}, for reciprocating country pairs (green) or non-reciprocating country pairs (red) and error bars represent the standard error. Probabilities (y-axis) have been shifted according to the aggregate probabilities of cooperation or conflict, respectively, across the entire ICEWS dataset.

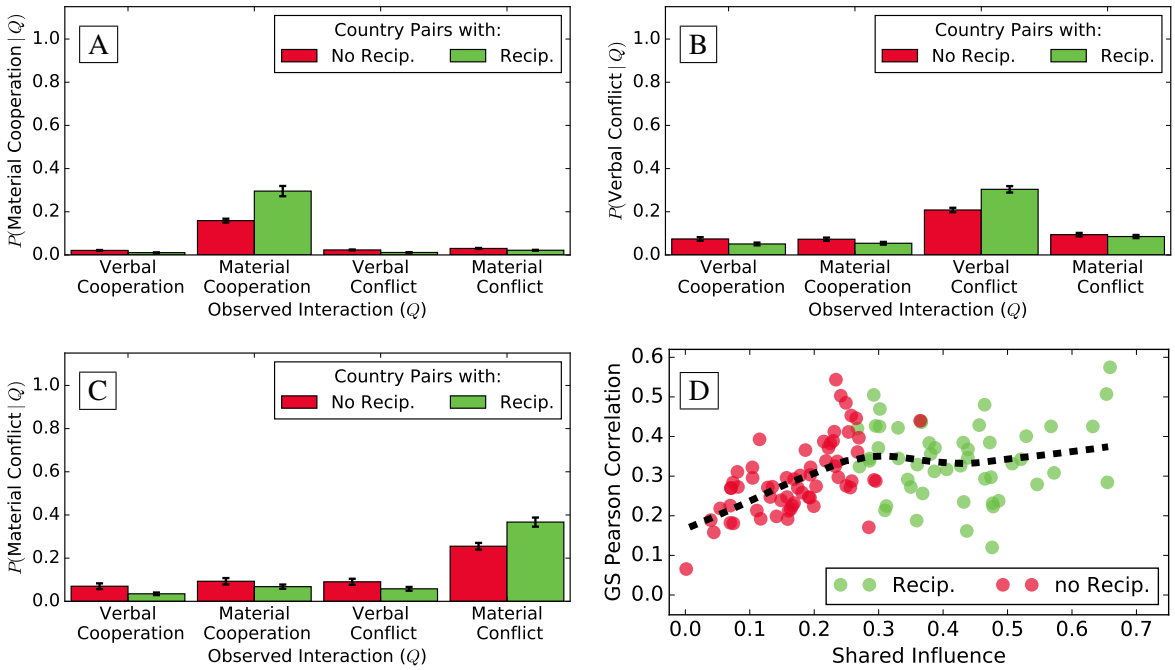

Figure 3: CCM reciprocity is indicative of correlated attitudes and mirroring of specific interaction types. **(A)-(C)** Given an observed interaction type (x-axis, denoted Q) between a country pair with CCM reciprocity (green) or without (red), we plot the probability (y-axis) of (A) Material Cooperation, (B) Verbal Conflict, and (C) Material Conflict in the day following the interaction (see S6 for additional time windows). **(D)** State-pairs with higher shared influence (i.e. $(CCM(A,B) + CCM(B,A))/2$, x-axis) have increasingly correlated attitudes towards each other (y-axis). Marker colors indicate if the country pair has reciprocity (i.e. $CCM(A,B) \geq 0.25$ & $CCM(B,A) \geq 0.25$, green) or not (red). The black dashed line represents a LOWESS regression fit.