aptara | NYAS | nyas12142-1585004 | Dispatch: April 30, 2013 | CE:
Journal | MSP No. | No. of pages: 7 | PE: Allie/Kelly
Ann. N.Y. Acad. Sci. ISSN 0077-8923

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: *Social Brain Conference: Neuroethics*

Trust and cooperation: a new experimental approach

Cristina Acedo and Antoni Gomila

Department of Psychology, University of the Balearic Islands, Palma de Mallorca, Spain

Q1 Address for correspondence: Antoni Gomila, Department of Psychology, University of the Balearic Islands, Campus Carretera Valldemossa Km 7.5, 07122 Palma de Mallorca, Spain. toni.gomila@uib.cat

Several theories within different disciplines emphasize the role of trust in fostering cooperation in human social life. Despite differences, the core of these notions of trust is affectively motivated loyalty, which makes the individuals feel mutually committed and willing to accept vulnerability because of positive expectations about each other's behavior. In evolutionary game theory and experimental economics, the notion of trust is much simpler: it is an expectation about another's behavior, a kind of wager, in which the sense of mutual commitments and vulnerability is completely absent. In order to extend the paradigm of trust games typical in those fields to explore the fuller sense of trust relationships, we have developed a new experimental design, in which an iterated prisoner dilemma is played by participants who do or do not hold a trusting personal relationship, while anonymity is preserved. We present here the results of our two pilot studies, which indicate the relevance of personal trust in fostering cooperation and suggest the influence of the structure of social networks on the degree of cooperation achieved.

Keywords: trust; cooperation; iterated prisoner dilemma; mutual commitment; normative expectations; social network

Since Mauss's *Essay on the Gift*[1] and Arrow's *The Limits of Organization*,[2] the notion that trust fosters human cooperation has become a kind of truism. However, it is a truism that has not been easily accommodated within the framework of rational decision theory of collective action. Several proposals from different disciplines have underlined the role of trust in social life as a missed dimension in the mainstream approach to the social sciences. From the theory of social capital in sociology,[3–6] to Ostrom's theory of social self-regulation in institutional economics and political science,[7,8] trust has received much attention as a key social factor that cements social groups, fosters cooperation, deters free riding, and averts social dilemmas.[9,10]

All these theories share an understanding of trust as an extrafamiliar, affective relationship that grounds mutual normative expectations and reciprocal commitments. These norms may be explicitly formulated, but are often just implicitly assumed. In any case, they are as different from contracts as they are from rational calculations. When engaging in a relationship of this kind, one puts oneself at risk of being exploited while believing that this risk is minimal or nonexistent, and one is aware that the partner runs a risk but expects not to be exploited, because of an affective binding relationship. Social psychologists have studied these kinds of relationships and tried to understand how trust is built, kept, and restored in case of a breach.[11] Social psychology also distinguishes between this central phenomenon of *particular* or personal trust,[12] and *generalized* or impersonal trust, that is, a general disposition regardless of effective relationships of trust related to personality differences, contextual and institutional factors, and previous experience.[13,14] Generalized trust, for instance, has been shown to be sensitive to the in-group/out-group dimension: people naturally trust members of the same social group more than members of other groups.[15–18] There has also been interest in studying the means by which individuals identify others as being trustworthy.[19]

This trend in the social sciences contrasts with the way trust has been dealt with by game theory, where approaches that use it as the basic tool, such as experimental economics and evolutionary theory, take for granted a self-interested view of the participating agents.[20,21] Within this framework, the

doi: 10.1111/nyas.12142

notion of trust is rather different: it is seen as a sort of prediction, an expectation about others' behavior, which can be viewed as a proxy for the social psychological notion of generalized trust.[22,23] So-called trust games, in which any contribution to the common pool generates further resources that are equally distributed, do not, in fact, involve relationships of personal trust among agents, but rather a kind of bet on what others will do. Therefore, the normative and affective dimensions of trust, and the reciprocal commitments, are absent in such games.

Trust games have shown that agents exhibit a disposition to cooperate, which is contrary to the predictions of rational decision theory. Between 40% and 60% of subjects contribute assets to the provision of the public good in a one-shot public good game,[24,25] and about the same percentage of participants contribute to the common pool in the first round of a finitely repeated prisoner's dilemma. The rate of contribution decays over time, but never reaches the predicted zero level.[26] Of course, personal trust may develop out of a series of mutually benefiting interactions, so that what could begin as a "tit-for-tat" strategy that is reciprocated may end up fostering real trust.[27,28] However, given the anonymous way such games are played, this possibility is, in fact, excluded in principle from consideration.

To experimentally study the role of the core, personal notion of trust in cooperation, we have tried to develop a new experimental design to test the influence of personal, particular trust relationships in social cooperation. We also want to investigate whether the structure of trust relationships within a group has an influence on the level of cooperation achieved.[29,30]

In this report, we present this experimental design and the result of two pilot studies in which it has been implemented. In our procedure, we ask each participant belonging to the same group to play an iterated prisoner's dilemma with either an anonymous individual from within her network of trust within the group or with someone from the group but outside of her network of trust. This second condition amounts to a control condition, where cooperation is expected to be similar to standard levels found in public goods/trust games. In addition, we manipulate whether the participants know how many rounds of the game they are to play. In the first study, participants knew how many iterations the game was going to have (a finite supergame), whereas in the second they did not, to approximate an infinite supergame.[31] In this manner, we may distinguish between cases where cooperation is the outcome of a self-interested strategy that has reached the cooperation equilibrium, and cases where cooperation is, in fact, the outcome of trust. Although agents initially deciding to cooperate on the grounds of a rational strategy, such as tit-for-tat, can keep cooperating in an infinite supergame (if that equilibrium is reached, even if cooperation may decrease due to the discount rate of the players), they are expected to defect at least on the last round of the game, if they know that it is the final round,[32,33] while trusting agents will keep cooperating.

To measure the influence of personal trust on cooperation, we developed four indicators of cooperation: (1) the proportion of participants selecting cooperation as their first choice; (2) the proportion of fully cooperative participants; (3) the proportion of full mutual cooperation along the game; and (4) the difference in cooperation choices between the first and third rounds of the game. The results indicate that more cooperation, in any way that it could be measured, took place in the personal trusting condition (TC) than in the nontrusting condition (NTC). The results reached statistical significance in just one of the pilot studies, the one in which the participants knew how many rounds of the game they would be playing. As expected, it was also the group with a significantly higher measure of personal versus generalized trust. Interestingly, this higher degree of personal trust correlates with a different structure of trust relationships within the group.

Method

Participants

Two studies were conducted: one with PhD students and postdocs from a research center at the University of Cambridge (UCAM); and another study with first-year education students from the University of the Balearic Islands (UIB). In both studies, a whole group was recruited.

Study 1: UCAM. This group comprised 13 participants, seven males and six females, aged 24–41 years old ($M = 30.46$, SD $= 5.06$), of different nationalities; four were Catholics, one was Buddhist, and eight were of no religious affiliation. Most were

unmarried and monthly income was variable: three below €500, four between €500 and €1000, and six between €1000 and €3000 euros. Their participation was completely voluntary.

Study 2: UIB. This group comprised 31 participants, 9 males and 22 females, aged 18–38 years old ($M = 20.35$, SD $= 5.03$), of Spanish nationality, 41% Catholics and 59% declared nonbelievers; mostly students, unmarried, and with monthly expenses below €500; just 23% of them declared to spending €500–1000 per month. They participated in exchange for course credit, and an additional prize depending on how many points they gained. At the time of participation, they did not know either what the prizes were or how many points they would need to obtain them. It was later revealed to them what they had gained: from drinks to a memory stick.

Procedure

Questionnaires. In both studies, participants first filled in first two questionnaires prepared by us. One questionnaire explored generalized trust, to control for the possible influence of this factor, involving several subscales, including institutional trust, personality traits (openness, extraversion, responsibility, neuroticism, and agreeableness), previous experiences of disappointment, etc., that may influence whether the individual would cooperate or defect[34] (see Supporting Information). The second questionnaire gauged personal trust: each participant had to write the names of three people in the group they trusted, and answer a set of questions about each relationship (see Supporting Information). From these questionnaires, we obtained measures of generalized and personal trust for each group. Participants were asked not to discuss their answers to the questionnaires. The questionnaires allowed us to give each participant scores of personal and generalized trust on a scale from 0 to 100.

Based on their responses to the second questionnaire, we assigned each of the participants to a different partner, for each condition, and appointed them for a second session, a couple of weeks later, in which there were to play the iterated prisoner's dilemma.

Iterated prisoner's dilemma. Each participant twice performed an iterated prisoner's dilemma,

P1	P1	P2	P2
C	3	3	C
D	6		C
C		6	D
D	1	1	D

Figure 1. **The prisonner's dilemma payoff matrix: C means cooperate and D means deceive. P1 is participant 1 and P2 is participant 2.**

with three repetitions. In the first condition, each participant played with one person she said she trusted, without knowing which one in particular, to maintain anonymity ("trust circle" condition: TC); in the second condition, each participant played with someone from the same group, but who was not mentioned as particularly trusted (NTC). Both conditions were presented in a counterbalanced order, and all participants did both.

In the UCAM study, participants knew that the game would be repeated three times. In the UIB study, they did not know how many rounds would be played. In both studies, the participants played at the same time but in different places, communicating through the researchers, to maintain anonymity.

The prisoner's dilemma was presented first to participants in a training phase, where they had to make five choices in a row as practice, to make sure they understood the pay-off matrix (Fig. 1).

After that, they were informed of the assigned condition (TC or NTC), and depending on the group, the number of rounds to play; then they were asked to decide whether to cooperate or deceive, without knowing the partner's decision. In the second and third rounds, the participants were informed of the partner's choice in the previous round, and the points obtained. Both participants had the same information and were under the same conditions.

At the end, the points obtained in each decision were added. This concluded the UCAM study. In the UIB study, the points obtained were converted to corresponding prizes a week later.

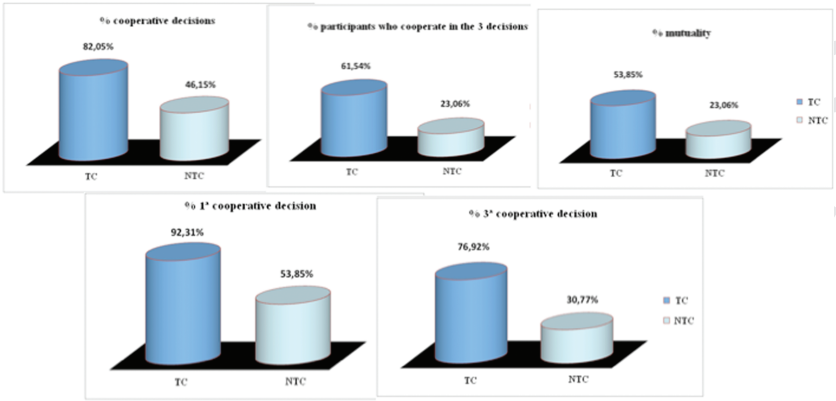


Figure 2. The results of the UCAM study.

COLOUR ONLINE, B&W IN PRINT

Results

The Influence of Personal Trust on Cooperation

To measure the level of cooperation, we have developed four indicators: (1) the proportion of cooperation choices; (2) the proportion of participants who cooperate in the three decisions; (3) the proportion of mutuality, that is, when both players always cooperate; and finally, (4) the contrast between the cooperative decisions in the first decision and in the third. We compared these data in both studies and under both conditions.

Study 1: UCAM. The results for the UCAM group are shown in Figure 2. One-tailed McNemar tests demonstrated that there was a significantly higher proportion of cooperation in the trust condition than in the other condition, for each dependent variable. Thus, the percentage of cooperation was 82.05% in TC, and 46.15% in NTC ($(1, n = 13) = 4$, $P < 0.02$); the percentage of participants who always cooperated was 61.54% in TC and 23.06% in NTC ($(1, n = 13) = 5$, $P < 0.05$); the instances of mutuality were 53.85% in TC and 23.06% in NTC ($(1, n = 13) = 4$, $P < 0.05$); and in the TC cooperation shifted from 92.31% in the first round to 76.92% in the third, while in the NTC, cooperation shifted from 53.85% to 30.77% (for the first round, $(1, n = 13) = 5$, $P < 0.05$); for the third, $(1, n = 13) = 4.5$, $P < 0.05$). Note that this group was told that they would play the IPD three times, and there were no prices delivered at the end.

The UCAM group scored 61.5 in the generalized trust questionnaire, and 81.3 in the personal trust questionnaire, which indicates a significantly higher level of personal trust among its members.

Study 2: UIB. The results for the UIB group are shown in Figure 3. One-tailed McNemar tests demonstrated that there was not a significantly higher proportion of cooperation in the trust condition than in the other condition, for each dependent variable in this group. Thus, the percentage of cooperation was 59.14% in TC and 47.31% in NTC ($(1, n = 31) = 0.305$, $P < 0.25$); the percentage of participants who always cooperated was 25.81% in TC and 12.9% in NTC ($(1, n = 31) = 2.66$, $P < 0.2$); the instances of mutuality were 10% in TC and 0 in NTC ($(1, n = 31) = 3$, $P < 0.1$); and in the TC cooperation shifted from 74.19% in the first round to 41.94% in the third, while in the NTC cooperation shifted from 64.52% to 38.71% (first round, $(1, n = 31) = 0.692$, $P < 0.5$; third round, $(1, n = 31) = 0.066$, $P < 0.5$). In this case, the participants were not told how many times they would play, and received prices in exchange for the points obtained in the game.

This group scored 71 in the "generalized" trust questionnaire, and 69.04 in the "personal" trust questionnaire, which is not a statistically significant difference.

The Structure of the Group

Given the procedure, we were also able to represent the trust relationships within the groups in each study (Fig. 4). Each member is represented as a node in a social network, and the size of the node represents social relevance (the number of people that trust this person). In these networks, the groups

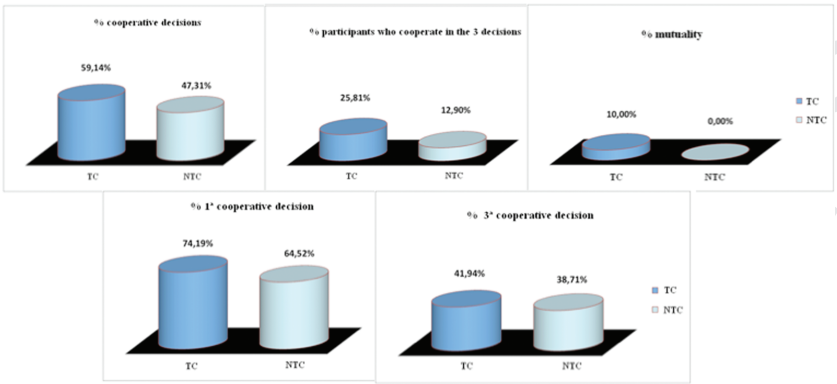


Figure 3. The results of the UIB study.

exhibit different patterns: while the UCAM network has a spider web form, revealing a high degree of cohesion and integration dependent upon close-knit groups of central individuals, the UIB network exhibits a more scattered distribution of trust. There are even isolated subgroups (cliques). The network appears as weakly integrated, crucially dependent upon a reduced number of individuals. In this network, we find greater number of coupled relationships but less cohesion.

Discussion and Conclusion

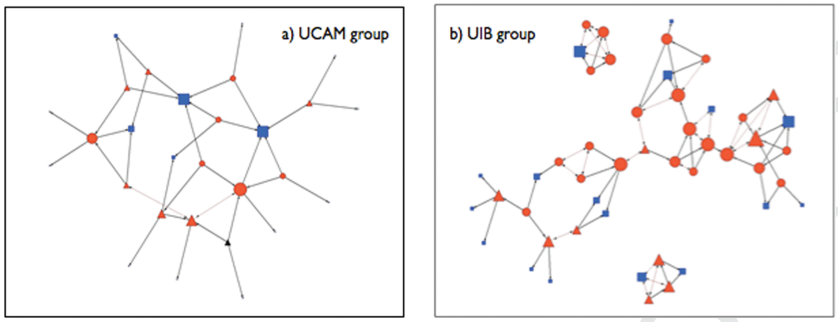
The pilot studies provide support to the notion that personal trust fosters cooperation, and offer a promising new experimental methodology to study it. As expected, the different measures developed clearly indicate a greater degree of cooperation when playing with a trusted partner than when playing with somebody else (even if from the same group). In the latter case, we found levels of cooperation similar to those found in standard trust/public goods games, those which block the possibility of personal relationships playing a role in the interaction. Personal trust fosters higher levels of cooperation and reciprocity: The results suggest that the higher the strength of the trust bond, the greater the cooperative attitude.

Cooperation was also higher in the trust condition in the UIB study, but the differences failed to reach statistical significance, probably because of the lower degree of personal trust, in comparison with the Cambridge group, while their respective measures of generalized trust were similar. The fact that the UIB group that volunteered to participate was one of first-year undergraduates may explain their lower level of personal trust, and hence a smaller difference between personal and generalized trust. Interestingly, this difference in particular trust between the groups seems to be connected to the different topologies of their respective social networks, but in a way that has yet to be explored.

Equally, as expected, trust helps overcome the well-documented trend of declining cooperation as the end of the game approaches (or because of applying a discount rate). In the Cambridge group, where personal trust is high, mutuality was also remarkably high, over 50%, even when the number of iterations was known, and during the last round of the game. The UIB group, with lower personal trust, showed the standard decline in cooperation, reaching a similar level in both conditions.

However, these studies are not enough to prove the point, even if one of the studies reached statistical significance: the procedure is now being applied to bigger groups, better chosen in terms of group level of personal trust. But these studies have been useful in devising the procedure, and finding ways to improve it. Personal trust measurement also needs more stringent criteria. On the other hand, it might be that playing for just pretend incentives, as in the Cambridge study, makes cooperation easier.

In this regard, the most intriguing result of these pilot studies concerns the role of the topology of the social network in fostering cooperation, beyond the personal relationships between each pair of members.[35] Personal trust, in addition to reputation, might account for the emergence of indirect reciprocity in a group,[36] through the impact it may have upon social cohesion and integration. This raises the possibility that the higher levels of



COLOUR ONLINE, B&W IN PRINT

Figure 4. **The group networks of both studies. Circular nodes represent female participants and triangular nodes represent male participants. Red nodes refer to the study participants and square and blue nodes refer to mentioned trustee individuals not involved in the study. Size node refers to nodal degree.**

cooperation in the UCAM group reflect not only a higher degree of personal trust but also to the structure of the social network. Group size (13 vs. 31) may also play a role in this regard, as well as the number of nominated trustees in the questionnaire. This is something that deserves further investigation, by means of an agent-based evolutionary model.

Overall, we found few people defecting all the time, and few people cooperating no matter what. The most common strategy was conditional cooperation, where trust increases the possibilities to cooperate, and cooperation, in turn, reinforces trust.[37] But there also appears to be two types of conditional cooperators in anonymous interactions: collaborators, who cooperate if their partner also cooperates, and deceivers, who cooperate to first encourage partner cooperation and then cheat on them. Finding ways to better infer agents' strategies from their behavioral choices is needed, just as experimental research needs to be combined with anthropological fieldwork,[38] to further our understanding of trust in the real world.

Acknowledgments

This research has been funded by Research Project FFI2010-20759 (Spanish Ministry of Science and Innovation). CA is supported by a PhD scholarship from the Spanish Ministry of Education (AP2009-2423). We thank Lluis Barceló and Julia Christensen for their continuous support, and are especially grateful to Philip Marlowe, for making possible CA's short stay in Cambridge, which was instrumental in carrying out this work, and providing guidance. We also want to express our gratitude to the reviewers, for their helpful comments and suggestions.

Supporting Information

Additional supporting information may be found in the online version of this paper.

Conflicts of Interest

The authors declare no conflicts of interest. Q3

References

1. Mauss, M. 1925. Essai sur le don. Forme et raison de l'échange dans les sociétés archaïques. L'Année Sociologique. (English translation by Ian Cunnison, The gift: the form and reason of exchange in archaic societies, Cohen & West, 1966; this translation is in the public dominion and can be found here: http://archive.org/details/ giftformsfunctio00maus). Q4
2. Arrow, K.J. 1974. *The Limits of Organization*. New York: W. W. Norton.
3. Coleman, J. 1988. Social capital in the creation of human capital. *Am. J. Sociol.* **94:** S95–S120.
4. Coleman, J. 1990. *Foundations of Social Theory*. Cambridge, Massachusetts: Belknap Press.
5. Fukuyama, F. 1995. *Trust: The Social Virtues and Creation of Prosperity*. New York: Free Press.
6. Putnam, R. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.
7. Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.
8. Poteete, A., M. Janssen & E. Ostrom. 2010. *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*. Princeton: Princeton University Press.
9. Rothstein, B. 2005. *Social Traps and the Problem of Trust: Theories of Institutional Design*. Cambridge, UK: Cambridge University Press.
10. Ostrom, E. 2009. Beyond markets and states: polycentric governance of complex economic systems. Nobel Prize Lecture, 8/xii/2009. Available at http://www.nobelprize.org/ nobel˙prizes/economics/laureates/2009/ostrom-lecture.html.
11. Lount, R.B., C.B. Zhong, N. Sivanathan & J.K. Murnigham. 2008. Getting off on the wrong foot: the timing of a breach

and the restoration of trust. *Pers. Soc. Psychol. B* **34:** 1601–1612.
12. Conviser, R.H. 1973. Toward a theory of interpersonal trust. *Pac. Sociol. Rev.* **16:** 377–399.
13. Bjørnskov, C. 2006. Determinants of generalized trust: a cross-country comparison. *Public Choice* **130:** 1–21.
14. Freitag, M. & R. Traunmüller. 2009. Spheres of trust: an empirical analysis of the foundations of particularised and generalised trust. *Eur. J. Polit. Res.* **48:** 782–803.
15. Chen, R. & Y. Yan Chen. 2011. The potential of social identity for equilibrium selection. *Am. Econ. Rev.* **101:** 2562–2589.
16. Chen, Y. & S.X. Li. 2009. Group identity and social preferences. *Am. Econ. Rev.* **99:** 431–457.
17. Eckel, C.C & P.J. Grossman. 2005. Managing diversity by creating team identity. *J. Econ. Behav. Organ.* **58:** 371–392.
18. Brañas-Garza, P., M.A. Durán & M.P. Espinosa. 2012. Favouring Friends. *Bull. Econ. Res.* **64:** 172–178.
19. Janssen, M. 2008. Evolution of cooperation in a one-shot Prisoner's dilemma based on recognition of trustworthy and untrustworthy agents. *J. Econ. Behav. Organ.* **65:** 458–471.
20. Camerer, C.F. 2002. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton: Princeton University Press.
21. Ostrom, E. & J. Walker, Eds. 2003. *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research.* New York: Russell Sage Foundation.
22. Deutsch, M. 1958. Trust and suspicion. *J. Conflict Resol.* **2:** 265–279.
23. Deutsch, M. 1973. *The Resolution of Conflict: Constructive and Destructive Processes.* New Haven: Yale University Press.
24. Davis, D. & Ch. Holt. 1993. *Experimental Economics.* Princeton: Princeton University Press.
25. Offerman, T. 1997. *Beliefs and Decision Rules in Public Goods Games: Theory and Experiments.* Dordrecht, Holland: Kluwer.
26. Isaac, R.M. & J.M. Walker. 1988. Group size effects in public goods provision: the voluntary contributions mechanism. *Q. J. Econ.* **103:** 179–199.
27. Yamagashi, T., S. Kanazawa, R. Mashima & S. Terai. 2005. Separating trust from cooperation in a dynamic relationship: prisoner's dilemma with variable dependence. *Ration. Soc.* **17:** 275–308.
28. Poteete, A., D. Janssen & E. Ostrom. 2010. *Working Together: Collective Action, the Commons, and Multiple Methods in Practice.* Princeton: Princeton University Press.
29. Curry, O. & R.I.M. Dunbar. 2011. Altruism in networks: the effects of connections. *Biol. Lett.* **7:** 651–653.
30. Kovářík, J., P. Brañas-Garza, R. Cobo-Reyes, *et al.* 2012. Prosocial norms and degree heterogeneity in social networks. *Physica A* **391:** 849–853.
31. Dal Bó, P. & G. Fréchette. 2011. The evolution of cooperation in infinitely repeated games: experimental evidence. *Am. Econ. Rev.* **101:** 411–429.
32. Murnighan, J.K. 1981. Defectors, vulnerabiity, and relative power: some causes and effects of leaving a stable coalition. *Hum. Relat.* **34:** 589–609.
33. Rubenstein, A. 1982. Perfect equilibrium in a bargaining model. *Econometrica* **50:** 97–109.
34. Acedo, C. & A. Gomila. 2012. Trust, social capital, and the evolution of human sociality. In *Proceedings of the Asian Conference on Psychology and the Behavioral Sciences*, Osaka, pp. 492–506.
35. Cook, KS & R. Hardin. 2001. "Norms of cooperativeness and networks of trust." In *Social Norms.* M. Hechter & K.-D. Opp, Eds. New York: Russell Sage Foundation.
36. Nowak, MA & K. Sigmun. 2005. Evolution of indirect reciprocity. *Nature* **437:** 1291–1298.
37. Six, F. 2005. *The Trouble with Trust. The Dynamics of Interpersonal Trust Building.* Edward Elgar Publishing Limited. **Q5**
38. Henrich, J., R. Boyd, S. Bowles, et al., Eds. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies.* New York: Oxford University Press.

Queries

Q1 Author: Please check details of the corresponding author as typeset for correctness.
Q2 Author: Figures 2 and 3 are in poor quality. Please provide better quality Figures.
Q3 Author: Please check the "Conflicts of Interest" statement for correctness.
Q4 Author: Please provide the last accessed date for the link provided in Refs. [1 and 10], also check Ref. [1] as typeset for correctness.
Q5 Author: Please provide the location of the publisher in Ref. 37.