Toward an experimental exploration of the complexity of human social interactions

Redouan Bshary[a,1] and Nichola J. Raihani[b,1]

[a]*Institute of Biology, University of Neuchâtel, 2000 Neuchâtel, Switzerland; and* [b]*Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, United Kingdom*

Few colleagues would debate that the complexity of human cooperation is unmatched in the animal kingdom. There are strong arguments that the complexity of our social life is tightly linked to the evolution of our large brain, particularly the neocortex (1). However, this view has only partly been tested in experiments designed to test evolutionary scenarios. One reason for this shortcoming is that experimental research on the evolution of cooperation is strongly driven by theoretical models. Models, however, typically explore relatively simple scenarios to allow for analytical solutions. As a consequence, the models may fail to capture the full complexity of human cooperative interactions. To appreciate and understand the sophistication of human social behavior, it is therefore important to design experiments that go beyond the paths explored by theory. Such experiments, in turn, may provide inspiration for new models. An empirical study in PNAS uses this approach (2).

The authors explored cooperative behavior in a setting that combined contribution to a public good with the potential for partner choice. A bystander observed a 15-round public goods game between four players, knowing that (s)he would replace one of these four in the second game. Participation yielded larger payoffs than being a bystander. The bystander could actively exclude one player from the second game by paying a few monetary units or could opt for a random allocation of partners, which was free. Players were informed about the bystander's option before the game started, and hence had to incorporate tradeoffs between immediate and future gains. Further complexity was introduced in several ways. First, the information available to bystanders was incomplete: They could only observe two players per round. Second, for a small fee per round, players could pay to conceal their contribution from the bystanders. Third, bystanders could pay to hide their observation of a specific player. In some experimental groups, rounds of the public goods game were alternated with an indirect reciprocity game, whereas in others, the public goods game was alternated with the possibility of punishing coplayers.

Why Complex Experiments Are Necessary

The complexity of this experimental setup goes way beyond the limitations that evolutionary game theory imposes on explicit experimental testing of models. Why are such complex experiments important? As it stands, humans being confronted with the basic cognitive requirements of the necessarily simple assumptions of evolutionary cooperation models may correspond to other animals being exposed to a Skinner box: Artificially simple experimental scenarios prevent humans from showing the sophistication of their social behavior (3). On the other hand, these simple scenarios may push other animals to their limits in terms of cooperative behavior. The risk with comparing human and animal behavior in artificially simple scenarios is that we may falsely invoke similarity in the underlying cognitive processes between the study species and humans in terms of the strategies used to solve social dilemmas. In reality, the partner control mechanisms used by humans and animals to sustain cooperation may be similar with respect to game theoretical concepts but may rest on very different cognitive processes.

> **Subjects focused on excluding cheaters rather than on ensuring that the most cooperative players remained to interact with them.**

Consider the famous example of indirect reciprocity based on image scoring. Alexander's proposal (4) that many examples of humans helping others can be explained with the associated gain in reputation that, in turn, increases the likelihood of receiving help when needed has been formalized by several authors (5, 6). An explicit experiment based on the model by Nowak and Sigmund (5) revealed that humans help people who they observed helping others and that helpful individuals therefore earned more money than nonhelpers (7). Furthermore, gains in reputation may even promote stable contributions to public goods (8, 9), and individuals are more helpful if they perceive cues that they are being watched (10, 11). The human concern for reputation may be predicated on an awareness of how bystanders' perceptions affect their beliefs. However, nonhuman animals may demonstrate similar behavioral patterns in interactions with others without the need to invoke such complex cognitive processes. Indeed, the key elements necessary for this form of cooperation are not by themselves particularly cognitively demanding, and they certainly exist in other animals. Many animals interact in a communication network and pay attention to interactions between third parties (12). As a consequence of such eavesdropping, individuals may alter current (and future) behavior. Although such "audience effects" have been mainly studied in the context of competition (12), it has been demonstrated that cleaner wrasses, *Labroides dimidiatus*, behave more cooperatively toward client reef fish when they are observed by other clients (13). Thus, we predict that it will only be a matter of time until interaction patterns are described in a nonhuman animal that fit indirect reciprocity via image scoring. Similar arguments can be made for other important concepts for human cooperation, such as third-party punishment (14) and cultural group selection (15). Social learning, the basic aspect of cultural group selection, has been demonstrated in a great variety of animals (16), and sophisticated decision rules regarding who and which behaviors to copy have been both postulated and documented in nonhuman animals (16, 17). Third-party punishment, where a bystander punishes an individual that cheated its partner, has also been documented in a primitive form in fish (18). The paper by Rochenbach and Milinski (2) captures some of the complexity of real-life interactions in humans, and hence allows the authors to investigate whether subjects use more sophisticated decision rules during

Author contributions: R.B. and N.R. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 18307.

[1]To whom correspondence may be addressed. E-mail: redouan.bshary@unine.ch or nicholaraihani@gmail.com.

cooperative interactions than evolutionary models would propose.

The Main Results

The authors found sophisticated dynamic strategies on either side. Actors paid to conceal their low contributions and failure to cooperate in the indirect reciprocity game. Actors also paid to conceal their behavior when they dished out severe punishment to cheats. Bystanders, on the other hand, often paid to observe covertly players who had displayed high contributions during open observations. When choosing partners for interactions, bystanders excluded low contributors as social partners but did not use punishment behavior for their decision. Given the complexity and the novelty of the game, the results yield few answers and raise many questions for both empirical and theoretical future research. Here, we present a few that we found particularly intriguing. First, we note that the main criterion used by bystanders to exclude a player was the player's low contribution relative to the second lowest contributor. In contrast, the presence of a particularly cooperative player did not lead to the active exclusion of a less cooperative player, even though the lottery process would remove such a highly cooperative player with $P = 0.25$. Thus, subjects focused on excluding cheaters rather than on ensuring that the most cooperative players remained to interact with them. Another puzzling result was the apparent correlation between open and hidden contributions for individual players: Bystanders managed to exclude the overall lowest contributors despite the incomplete information they had. We would have expected that particularly strategic individuals would contribute most when observed and least when their contributions were concealed, causing mismatches between observed and average contributions.

Perhaps the most interesting results concern the treatment with punishment. First, it appears that the presence of this control mechanism causes a reduction in the variance between individual contributions. This effect undermines the need for the bystander to exclude an uncooperative player actively. In only 3 of 16 groups with punishment did the bystander pay to exclude a potential coplayer. This contrasts markedly with the other two conditions: six of eight bystanders paid to remove one player in the simple public goods condition, and four of eight made this active choice in the public goods game that alternated with indirect reciprocity rounds. Thus, punishment interacts with partner choice in interesting ways. It is also puzzling that players paid to conceal strong punishment of coplayers from bystanders. This indicates that humans think that others perceive punishment as something negative. Such thinking contrasts with the actual decision-making process by the bystanders, who did not use punishment behavior for partner choice decisions. The data also contrast with the assumption of a model demonstrating that punishment could evolve because it yields social prestige in similar ways as helping does (19).

Perspectives

So, where to go from here? An important next step will be to look at the payoff consequences of decisions: At what point do bystanders obtain a net benefit from paying to exclude the lowest contributor? Do the gains associated with low contributions in the first public goods game compensate for the risk of being removed from the second game? In other words, what decision rules would be under positive selection in this game? An interesting additional treatment would be to allow players to conceal their concealing decisions. In this way, players do not admit guilt when they conceal their contributions, as was the case in the current study: Players hid cheating decisions, and bystanders could guess that cheating was the reason for hiding. This may have been the reason why hiding was rare (on average, in only 1 of 15 trials). In a follow-up study, bystanders could be told that a subset of the players was unavailable to view; some of these would have been the concealed concealers, and some could have been randomly allocated concealed players. This would allow us to determine whether players pay more for concealed concealing than for "open" concealing and whether bystanders manage to incorporate that additional uncertainty in their choices. As the authors conclude themselves, their study may have uncovered only the tip of the iceberg.

ACKNOWLEDGMENTS. R.B. is financed by the Swiss Science Foundation. N.R. is financed by a Royal Society University Research Fellowship.

1. Dunbar RIM (1993) Coevolution of neocortex size, group size and language in humans. *Behav Brain Sci* 16:681–735.
2. Rockenbach B, Milinski M (2011) To qualify as a social partner, humans hide severe punishment, though their observed cooperativeness is decisive. *Proc Natl Acad Sci USA* 108:18307–18312.
3. de Waal FBM (2001) *The Ape and the Sushi Master: Cultural Reflections of a Primatologist* (Basic Books, New York).
4. Alexander RD (1987) *The Biology of Moral Systems* (Aldine de Gruiter, New York).
5. Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393:573–577.
6. Leimar O, Hammerstein P (2001) Evolution of cooperation through indirect reciprocity. *Proc Biol Sci* 268: 745–753.
7. Wedekind C, Milinski M (2000) Cooperation through image scoring in humans. *Science* 288:850–852.
8. Milinski M, Semmann D, Krambeck HJ (2002) Reputation helps solve the 'tragedy of the commons'. *Nature* 415:424–426.
9. Tennie C, Frith U, Frith CD (2010) Reputation management in the age of the world-wide web. *Trends Cogn Sci* 14:482–488.
10. Haley KJ, Fessler DMT (2005) Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol Hum Behav* 26:245–256.
11. Bateson M, Nettle D, Roberts G (2006) Cues of being watched enhance cooperation in a real-world setting. *Biol Lett* 2:412–414.
12. McGregor PK, ed (2005) *Animal Communication Networks* (Cambridge Univ Press, Cambridge, UK).
13. Pinto AI, Oates J, Grutter AS, Bshary R (2011) Cleaner wrasses *Labroides dimidiatus* are more cooperative in the presence of an audience. *Curr Biol* 21: 1140–1144.
14. Fehr E, Gächter S (2002) Altruistic punishment in humans. *Nature* 415:137–140.
15. Boyd R, Gintis H, Bowles S, Richerson PJ (2003) The evolution of altruistic punishment. *Proc Natl Acad Sci USA* 100:3531–3535.
16. Whiten A, Hinde RA, Laland KN, Stringer CB (2011) Culture evolves. *Philos Trans R Soc Lond B Biol Sci* 366:938–948.
17. Kendal JR, Rendell L, Pike TW, Laland KN (2009) Nine-spined sticklebacks deploy a hill-climbing social learning strategy. *Behav Ecol* 20:238–244.
18. Raihani NJ, Grutter AS, Bshary R (2010) Punishers benefit from third-party punishment in fish. *Science* 327: 171.
19. dos Santos M, Rankin DJ, Wedekind C (2011) The evolution of punishment through reputation. *Proc Biol Sci* 278:371–377.