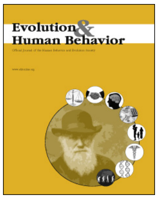
Evolution and Human Behavior xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org



Sustained cooperation by running away from bad behavior

Charles Efferson [a,*,1], Carlos P. Roca [b,*,1], Sonja Vogt [a], Dirk Helbing [c,d,*]

[a] *Department of Economics, University of Zurich*
[b] *Department of Chemical Engineering, Universitat Rovira i Virgili, Spain*
[c] *Chair of Sociology, In Particular of Modeling & Simulation, Swiss Federal Institute of Technology, Zurich*
[d] *Santa Fe Institute, USA*

ARTICLE INFO

Article history:
Initial receipt 11 November 2014
11 May 2015
Final revision received 15 May 2015
Available online xxxx

Keywords:
Evolution of cooperation
Positive assortment
Migration
Behavioral experiment

ABSTRACT

For cooperation to evolve, some mechanism must limit the rate at which cooperators are exposed to defectors. Only then can the advantages of mutual cooperation outweigh the costs of being exploited. Although researchers widely agree on this, they disagree intensely about which evolutionary mechanisms can explain the extraordinary cooperation exhibited by humans. Much of the controversy follows from disagreements about the informational regularity that allows cooperators to avoid defectors. Reliable information can allow cooperative individuals to avoid exploitation, but which mechanisms can sustain such a situation is a matter of considerable dispute. We conducted a behavioral experiment to see if cooperators could avoid defectors when provided with limited amounts of explicit information. We gave each participant the simple option to move away from her current neighborhood at any time. Participants were not identifiable as individuals, and they could not track each other's tendency to behave more or less cooperatively. More broadly, a participant had no information about the behavior she was likely to encounter if she moved, and so information about the risk of exploitation was extremely limited. Nonetheless, our results show that simply providing the option to move allowed cooperation to persist for a long period of time. Our results further show that movement, even though it involved considerable uncertainty, allowed would-be cooperators to assort positively and eliminate on average any individual payoff disadvantage associated with cooperation. This suggests that choosing to move, even under limited information, can completely reorganize the mix of selective forces relevant for the evolution of cooperation.

1. Introduction

Paradoxically, the evolution of human cooperation is an area of research characterized by both widespread agreement and seemingly interminable controversy. Regarding the widespread agreement, researchers generally accept that the evolution of cooperation requires some kind of informational regularity (Frank, 1998; Henrich, 2004; Nowak, 2006; van Veelen, 2009; Bowles & Gintis, 2011). In its most general form, this regularity can be thought of as a situation in which a cooperating individual is more likely than a defecting individual to interact with someone who cooperates. Such a regularity means that individuals who cooperate enjoy, with a relatively high probability, the benefits produced when others cooperate. This limits the risk of exploitation and produces mutual gains that can support the evolution of cooperation.

In spite of agreement on this point, a seemingly interminable controversy persists because researchers often disagree bitterly about what constitutes a reasonable explanation for the required informational regularity. Genetic relatedness due to common ancestry provides an uncontroversial explanation (Hamilton, 1964). The details of genetic transmission ensure that two actors with the same parents, as one example, will on average be more similar to each other than two individuals randomly selected from the population. This is an example of an informational regularity that can support the evolution of cooperation, and cooperation among kin is well established, well understood, and widely accepted.

Reciprocal strategies can also generate the required regularity. Such strategies share the feature that a focal individual conditions her decision to cooperate in some way on information about current or potential partners. Conditional cooperation can arise from simple rules based on direct experience with a partner (Axelrod, 1984) or indirect experience (Nowak & Sigmund, 1998), and they can also involve subtle considerations that assign good or bad standing to a partner (Leimar, 1997; Panchanathan & Boyd, 2004; Nowak & Sigmund, 2005). Though reciprocal strategies can evolve under the right circumstances, they can also be vulnerable in many ways. They are often not especially robust as group size increases (Boyd & Richerson, 1988), when agents make errors

---

* Corresponding authors.
E-mail addresses: charles.efferson@econ.uzh.ch (C. Efferson), carlosp.roca@urv.cat (C.P. Roca), dhelbing@ethz.ch (D. Helbing).

[1] Efferson and Roca contributed equally to this research.

http://dx.doi.org/10.1016/j.evolhumbehav.2015.05.003

Please cite this article as: Efferson, C., et al., Sustained cooperation by running away from bad behavior, *Evolution and Human Behavior* (2015), http://dx.doi.org/10.1016/j.evolhumbehav.2015.05.003

(Nowak & Sigmund, 1993; Henrich, 2004; Nowak & Sigmund, 2005), or when the space of admissible strategies is sufficiently complex (Boyd & Lorberbaum, 1987; van Veelen, García, Rand, & Nowak, 2012).

In any case, with respect to human behavior the larger puzzle is that humans routinely cooperate with genetically unrelated individuals in situations where reciprocal behavior is either not possible or cannot produce future benefits for the reciprocator (Camerer, 2003; Fehr & Fischbacher, 2003; Henrich et al., 2006; Raihani & Bshary, 2015). This simple observation leads to questions of the following sort. Do humans cooperate with unrelated individuals because they are responding to a kinship-based psychology that is occasionally misapplied, or do humans have prosocial motives that evolved in part for reasons unrelated to genetic relatedness (Fehr & Henrich, 2003)? One can ask analogous questions about cooperation in anonymous one-shot interactions and the evolution of reciprocal strategies (Burnham & Johnson, 2005; Haley & Fessler, 2005; Hagen & Hammerstein, 2006; Burnham, 2013). Answers to these questions vary, but the upshot is that at least some researchers have concluded that conventional evolutionary explanations based on kinship and reciprocity are not sufficient to explain human cooperation in its entirety (Henrich, 2004; Boyd & Richerson, 2005; Bowles & Gintis, 2011).

Alternative explanations have been offered, and shared group affiliation in a structured population is a controversial one (Williams, 1966; Soltis, Boyd, & Richerson, 1995; Boyd & Richerson, 2005; Bowles, 2009; Boyd, Richerson, & Henrich, 2011). If variation in cooperation levels occurs mostly between groups, then individuals in the same group tend to be similar. Group affiliation in such a case is an important source of information. Interacting with a group affiliate increases the probability that a cooperator interacts with a cooperator, and it decreases the probability that a defector interacts a cooperator. The problem with shared group affiliation as a source of information, however, is maintaining variation between groups. In a strictly genetic system at least (cf. Boyd et al., 2011), selection within groups and migration will often eliminate most of the differences between groups (Henrich, 2004). This destroys the information associated with shared group affiliation, and in the end a randomly selected member of one's own group will be extremely similar, in expectation, to a randomly selected member of any group. Put differently, relatedness within groups should often be low (cf. Bell, Richerson, & McElreath, 2009; Hill et al., 2011).

The above reasoning holds when migration is global and unsystematic. If cooperators ever find themselves in groups of their own, migration of this sort will largely destroy this kind of grouping. What if, however, movement between groups is a biased process? Partner choice is an especially clear mechanism for generating biases in group formation (Noë & Hammerstein, 1995; Roberts, 1998; Raihani & Bshary, 2015). Unlike reciprocal strategies, which condition behavior on a current partner's recent behavior or reputation, partner choice conditions one's willingness to interact socially on a partner's recent behavior or reputation. To give an illustrative but somewhat unrealistic example, assume that cooperators, and only cooperators, always know of some new secret place to rendezvous if they face too much free-riding in their current groups. This means they can consistently sequester themselves, even if their secret places are eventually discovered, they can interact primarily with their fellow cooperators, and they can consistently enjoy the gains from mutual cooperation. Presumably this mechanism would support the evolution of cooperation, but it requires that cooperators have privileged access to crucial information. As with green beards (Dawkins, 1976) and image scores (Nowak & Sigmund, 1998; Wedekind & Milinski, 2000), the informational regularity that allows cooperators to limit exploitation is like a gift from heaven.

As this example suggests, however, effective partner choice can involve distinct mechanisms (Noë & Hammerstein, 1995). Specifically, one can opt out of a current relationship because one knows it is bad (Eguíluz, Zimmermann, Cela-Conde, & San Miguel, 2005; McNamara, Barta, Fromhage, & Houston, 2008; Roca & Helbing, 2011; Bednarik, Fehl, & Semmann, 2014), one can opt in to a new relationship because of information suggesting the relationship might be good (Barclay & Willer, 2007), or one can do both (Rand, Arbesman, & Christakis, 2011; Wang, Suri, & Watts, 2012; Antonioni, Cacault, Lalive, & Tomassini, 2014). Opting in can lead to prosocial behavior because competition for partners can lead players to use current cooperation to signal that they would make good partners (Roberts, 1998; Barclay & Willer, 2007). For people looking to opt in to a new relationship, this mechanism requires that the current cooperation of a potential partner is a reliable and available source of information about the potential partner's future behavior. As always, this raises questions about where this reliable information comes from and how its integrity is maintained.

Opting out of an existing relationship is different. Opting out relies instead on one's recent personal experience. One does not need to know how a potential partner might behave. Rather one needs to know how current partners have been behaving. The informational requirements in this case are considerably less stringent than for opting in. Recent models suggest that opting out might work as a stand-alone mechanism (McNamara et al., 2008; Roca & Helbing, 2011), but we know little about how people actually use the option to leave others behind. In particular, we would like to know how effective is a situation in which players can opt out of their current relationships, but they have no information about the kinds of partners they will face after doing so. Put differently, how much cooperation results when players cannot run toward good behavior, but they can run away from bad behavior?

To find out, we conducted a behavioral experiment in which players were distributed on a lattice. In every period, each player had the option to move to a new location. After player movements, each player played a social dilemma game with all the players in her neighborhood. Crucially, information was very limited, and thus players faced considerable uncertainty when they moved. In particular, players were not identifiable as individuals, and their histories of play were not publicly available. Thus, a player who chose to move did not know with whom she would interact in her new location, nor whether her new partners were likely to cooperate. When a player moved, in short, she took a shot in the dark.

What a player did know was how she had fared in her past interactions. Specifically, a player knew how many partners she had recently played the game with, and she knew the payoffs she had received. Players could thus infer the total cooperation in their respective neighborhoods but little else. The key question is the following. When coupled with the option to migrate away from one's current neighborhood, can this kind of limited information allow players to avoid exploitation often enough to reduce or even eliminate the disadvantages of cooperating? If so, the result would show that non-random movement based on limited information could have been an important mechanism contributing to the evolution of human cooperation.

2. Experimental methods

Our experiment consisted of two treatments. To ease the exposition, we will first describe our main experimental treatment, the "choose location" treatment. Afterwards, we will describe the control, which we call the "assign location" treatment.

In the choose location treatment, each player was located somewhere on a 12 × 12 lattice. As explained below, sessions involved 31 or 35 participants, and as a result 20%–25% of available sites were occupied. Because of this low density, participants who wanted to move were not limited to recently vacated sites.

The lattice was on a torus, and thus no player was ever on an edge or corner. At the beginning of a period, each player saw the current state of the lattice (Fig. 1). She saw herself at the center of the lattice, and she saw which cells were occupied for the rest of the lattice. In addition, each player saw two neighborhoods of interest. She saw her eight-cell

Please cite this article as: Efferson, C., et al., Sustained cooperation by running away from bad behavior, *Evolution and Human Behavior* (2015), http://dx.doi.org/10.1016/j.evolhumbehav.2015.05.003