

Models inconsistent with altruism cannot explain the evolution of human cooperation

Kristian Ove R. Myrseth^{a,1} and Conny E. Wollbrant^{b,1}

The article “Intuition, deliberation, and the evolution of cooperation,” by Bear and Rand (1), uses game theoretic models to examine the role of intuition and deliberation in human cooperation. The premise is that dual processes characterize human social decision making: “(i) automatic, intuitive processes that are relatively effortless but inflexible; and (ii) controlled, deliberative processes that are relatively effortful but flexible” (1). The objective is to “provide a formal theoretical framework for considering the question of whether prosociality is intuitive or whether it requires self-control,” and the article concludes that “evolution never favors strategies for which deliberation increases cooperation” (1). However, the evolutionary model suffers from a serious shortcoming; it precludes the survival of altruistic individuals—thought to represent a major share of human populations (2). It is therefore not suitable for addressing whether human cooperative behavior is intuitive.

Although the model can account for prosocial behavior in one-shot interactions, such as one-shot prisoner’s dilemmas, the cooperation observed is rooted in self-interest and explained as a spillover “from settings where cooperative behavior can be payoff-maximizing.” Such strategic cooperation stands in contrast to intrinsically altruistic behavior, which is thought to represent a defining feature of human social interaction, across cultures (3). A case in point is the concept of strong reciprocity—which combines the altruistic propensity to reward others for cooperative, norm-abiding behavior with the predisposition to punish others for norm violations or noncooperative behavior, even when individually costly (3). Altruistic behavior often finds its

motivational source in prosocial emotions, among which the most important are “empathy”—the apprehension or comprehension of another’s emotional state—and “empathetic concern,” better known as sympathy (4). Another source is “warm glow,” or impure altruism—the individual is motivated by emotional rewards from acting prosocially (5). For example, subjective satisfaction—as well as neural activity in areas associated with reward processing, the caudate and the right nucleus accumbens—are amplified when individuals voluntarily make transfers to a charity (6).

The model put forth by Bear and Rand (1), however, precludes altruism in equilibrium. Individuals play either a one-shot or a repeated prisoner’s dilemma, and costly deliberation allows for revision of strategies, in case initial strategies are suboptimal for the particular game at hand. An altruist who plays the repeated prisoner’s dilemma would prefer to cooperate, but so would a self-interested individual. In the one-shot game, however, the altruist would also prefer to cooperate, although this strategy is suboptimal in material terms. This puts the altruist at a disadvantage—those who play the defect strategy, even if only occasionally, would do better in reproductive terms, implying extinction of altruistic preferences. Similar reasoning implies that a population of altruists would not survive the introduction of selfish players.

It is evident that the model by Bear and Rand (1) precludes a crucial stylized fact about human social decision making: Cooperation is not only strategically motivated—it is often altruistic. A meaningful model of the evolution of human cooperation must produce at least one equilibrium consistent with this fact.

- 1 Bear A, Rand DG (2016) Intuition, deliberation, and the evolution of cooperation. *Proc Natl Acad Sci USA* 113(4):936–941.
- 2 Fehr E, Schmidt KM (2006) The economics of fairness, reciprocity and altruism – Experimental evidence and new theories. *Handbook on the Economics of Giving, Reciprocity and Altruism*, eds Serge-Christophe K, Jean MY (Elsevier, Amsterdam), pp 615–691.
- 3 Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425(6960):785–791.
- 4 Jensen K, Vaish A, Schmidt MFH (2014) The emergence of human prosociality: Aligning with others through feelings, concerns, and norms. *Front Psychol* 5:822.
- 5 Andreoni J (1989) Giving with impure altruism: Applications to charity and Ricardian equivalence. *J Polit Econ* 97(6):1447–1458.
- 6 Harbaugh WT, Mayr U, Burghart DR (2007) Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316(5831):1622–1625.

^aSchool of Management, University of St. Andrews, Fife KY16 9RJ, United Kingdom; and ^bDepartment of Economics, School of Business, Economics and Law, University of Gothenburg, 405 30 Gothenburg, Sweden

Author contributions: K.O.R.M. and C.E.W. conceived the argument and wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence may be addressed. Email: kom@st-andrews.ac.uk or conny.wollbrant@economics.gu.se.