Evolution and Human Behavior 38 (2017) 421–427

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org

Original Article

A popular misapplication of evolutionary modeling to the study of human cooperation

Daniel Cownden [a,b], Kimmo Eriksson [b,c,*], Pontus Strimling [b,d]

[a] Institute of Infection, Immunity & Inflammation, University of Glasgow, 120 University Place, G12 8TA Glasgow, United Kingdom
[b] Centre for the Study of Cultural Evolution, Stockholm University, SE-10691 Stockholm, Sweden
[c] School of Education, Culture and Communication, Mälardalen University, Box 883, SE-72123 Västerås, Sweden
[d] Institute for Futures Studies, Box 591, SE-10131 Stockholm, Sweden

ARTICLE INFO

Article history:
Initial receipt 12 July 2016
Final revision received 9 January 2017

Keywords:
Altruism
Cultural evolution
Evolutionary game theory
Genetic evolution
Human cooperation
Phenotypic gambit

ABSTRACT

To examine the evolutionary basis of a behavior, an established approach (known as the phenotypic gambit) is to assume that the behavior is controlled by a single allele, the fitness effects of which are derived from a consideration of how the behavior interacts, via life-history, with other ecological factors. Here we contrast successful applications of this approach with several examples of an influential and superficially similar line of research on the evolutionary basis of human cooperation. A key difference is identified: in the latter line of research the focal behavior, cooperation, is abstractly defined in terms of immediate fitness costs and benefits. Selection is then assumed to act on strategies in an iterated social context for which fitness effects can be derived by aggregation of the abstractly defined immediate fitness effects over a lifetime. This approach creates a closed theoretical loop, rendering models incapable of making predictions or providing insight into the origin of human cooperation. We conclude with a discussion of how evolutionary approaches might be appropriately used in the study of human social behavior.

1. Introduction

Our aim in this paper is to offer a critique of a popular research paradigm. We shall refer to it as the puzzle paradigm as its use tends to be motivated by an assertion that human cooperation is an evolutionary puzzle. We use four influential papers as examples of the puzzle paradigm. As described in detail in Section 2, these papers are characterized by a two-step sequence of mathematical modeling: (1) cooperative behavior is modeled as an altruistic act in a giving game, in which payoffs are immediately transferable to fitness. Under this initial model it is puzzling that humans, unlike other comparable animals, seem to cooperate a great deal among non-kin, as the choice of an altruistic act in a giving game among non-kin should be an evolutionarily unsuccessful strategy. (2) The giving game is then embedded in an iterated social context. Selection is assumed to act on strategies in this iterated social context (rather than on behaviors in the underlying giving game). The fitness of such strategies is derived from the aggregation of the immediate fitness effects of the behaviors occurring in the iterated social context. This embedding is claimed to resolve the puzzle of human cooperation through a demonstration that strategies may involve altruistic acts and still be evolutionary successful.

We argue that the puzzle paradigm is implicitly employing a research strategy from evolutionary biology called the phenotypic gambit. In Section 3 we describe the phenotypic gambit and discuss some of the critical assumptions that justify its use. Two examples are given of careful application of the phenotypic gambit in evolutionary biology. In Section 4 we show how puzzle paradigm research deviates from a standard application of the gambit by contrasting it with these examples. Specifically we show how the puzzle paradigm practice of defining behaviors abstractly in terms of payoffs differs from the standard approach used by evolutionary theorists and how this leads to a closed theoretical loop, rendering puzzle paradigm models incapable of making predictions or providing insight into the origins of or the existent patterns of variation of human cooperation. In Section 5 we discuss how evolutionary approaches might be appropriately used in the study of human social behavior.

2. The puzzle paradigm

We have selected four influential papers (Boyd, Gintis, & Bowles, 2010; Boyd, Gintis, Bowles, & Richerson, 2003; Hauert, Traulsen, Brandt, Nowak, & Sigmund, 2007; Nowak & Sigmund, 2005) as canonical representatives of the puzzle paradigm. These papers, which

* Corresponding author at: Centre for the Study of Cultural Evolution, Stockholm University, SE-10691 Stockholm, Sweden.
E-mail addresses: dcownden@gmail.com (D. Cownden), kimmoe@gmail.com (K. Eriksson), pontusstrimling@gmail.com (P. Strimling).

http://dx.doi.org/10.1016/j.evolhumbehav.2017.01.004
1090-5138/

appeared in the highest ranked general science journals, study how cooperation in humans may have evolved in tandem with costly punishment mediated through a variety of mechanisms. Each paper begins by asserting or implying that human cooperation is an evolutionary puzzle:

> "Unlike any other species, humans cooperate with non-kin in large groups. This behavior is puzzling from an evolutionary perspective […]" (Boyd et al., 2003).

> "[…] human societies, are organized around altruistic, cooperative interactions. How can natural selection promote unselfish behavior?" (Nowak & Sigmund, 2005).

> "All human populations seem willing to use costly punishment to varying degrees, and their willingness to punish correlates with the propensity for altruistic contributions. This raises an evolutionary problem […]" (Hauert et al., 2007).

> "Humans are a uniquely cooperative species […] Recent theoretical studies provide an evolutionary explanation for such cooperative behavior […] There are two important problems with this explanation of human cooperation" (Boyd et al., 2010).

These quotes highlight the two distinct but related research questions the puzzle paradigm seeks to answer: 1) why are humans unique among species with regard to cooperation, and 2) how can cooperative behaviors (regardless of species) be selected for in an evolutionary process?

2.1. The first step: modeling cooperation as altruism in a giving game

In the puzzle paradigm cooperative behavior is operationalized as a behavior where an individual pays a personal fitness cost to bestow a fitness benefit upon others:

> "There are two behavioral types, contributors and defectors. Contributors incur a cost c to produce a total benefit b that is shared equally among group members. Defectors incur no costs and produce no benefits." (Boyd et al., 2003)

> "In the simplest model, the altruistic act consists in conferring a benefit b on the recipient at a cost c to the donor." (Nowak & Sigmund, 2005)

> "Those who participate can decide whether or not to contribute an investment at a cost c to themselves. All individual contributions are added up and multiplied with a factor $r > 1$. This amount is then divided equally among all participants of the public goods game." (Hauert et al., 2007)

> "Cooperation costs the cooperator c and benefits each member of the group b/n ($b > c > b/n$)." (Boyd et al., 2010)

Such a characterization, when considered in isolation, is synonymous with altruism. Thus, puzzle paradigm research models cooperative behavior as an altruistic act in some version of a giving game. Much confusion has arisen around the way different researchers use the terms altruism and cooperation, and this has been discussed at length by West, Griffin, and Gardner (2007). Whether or not a cooperative act should be considered truly altruistic in the sense of West et al. depends on the broader context of the behavior, and potentially its aggregate fitness consequences. See Bergmüller, Russell, Johnstone, and Bshary (2007) for a thoughtful discussion on immediate versus long run fitness consequences in the context of describing cooperative behaviors.

From these quotes it is clear that cooperation is defined in terms of the immediate payoffs in a giving game. Moreover, these payoffs are assumed to be in the currency of fitness. In other words, the strategy that incurs the largest net benefit will increase most in frequency. A "cooperative" strategy of paying a cost to give a benefit to others should therefore be outcompeted by a "non-cooperative" strategy. The emergence and persistence of high levels of cooperation under these conditions should not be possible.

2.2. The second step: embedding the giving game to resolve the puzzle

The puzzle paradigm's research goal is to resolve the proposed puzzle of human cooperation by showing that high levels of cooperation can emerge and persist under the additional assumption that the giving game is embedded in a particular iterated social context. This embedding varies from paper to paper and is rather lengthy to describe. For instance, the embedding of Boyd et al. (2003) amounts to the following four steps: (1) structuring the population into groups, with a certain amount of migration between groups; (2) adding a third strategy type that both cooperates and punishes non-cooperators in their group, at certain costs to themselves and the punished parties; (3) letting relative replication of strategies depend on the total payoffs obtained from the giving game and punishments; (4) letting groups meet each other in pairwise competition, with the probability of winning decided by the frequency of cooperation in each group, and with the winning group replacing the losers. Given this embedding it is then demonstrated that, as strategies replicate, cooperative behavior in the giving game and punishing behavior in the punishment game are both sustained at high levels.

The other three puzzle paradigm papers in our sample similarly embed the simple giving game in an iterated social context, each with some particular features: Nowak and Sigmund (2005) deal with reputation and indirect reciprocity, Hauert et al. (2007) with optional participation, and Boyd et al. (2010) with coordination and economies of scale.

Thus an implicit hypothesis of puzzle paradigm research is that human cooperation is enabled by the particular iterated social structures in which the giving game is embedded. The evolutionary puzzle initially posited is resolved to the extent that such social structures are unique to humans (and, of course, to the extent that human behavior in such social structures is indeed governed by the kind of strategies that are posited by the model).

3. The phenotypic gambit

A "gambit" is a chess opening in which a piece is sacrificed in order to gain what is hopefully a better position. The phenotypic gambit was perhaps first made by Fisher (1930) in an attempt to understand the evolution of sex ratios. The term phenotypic gambit was coined in Grafen (1984). We quote part of his definition:

> "The phenotypic gambit is to examine the evolutionary basis of a character as if the simplest possible genetic system controlled it: as if there were a haploid locus at which each distinct strategy was represented by a distinct allele, as if the payoff rule gave the number of offspring for each allele, and as if enough mutation occurred to allow each strategy to invade." (Grafen, 1984)

This quote emphasizes the simplification of genetic inheritance mechanisms, while perhaps deemphasizing the simplification of the complex and context dependent mapping from genotype to fitness via phenotype. We encourage readers to read Grafen's introduction of the term in its entirety, as well as other related accounts of when and why such a gambit is justified (Smith, 1978; Dawkins, 1999 pp. 30–54). Here we will do our best to paraphrase these accounts.