Cultural Transmission Can Explain the Evolution of Cooperative Behavior

Dor Cohen[1], Ohad Lewin-Epstein[2], Marcus W. Feldman[3], and Yoav Ram[a,*]

[1]School of Computer Science, Inetdisciplinary Center Herzliya, Herzliya, Israel
[2]School of Plant Sciences and Food Security, Tel Aviv University, Tel Aviv, Israel
[3]Department of Biology, Stanford University, Stanford, CA
[*]Corresponding author: yoav@yoavram.com

July 28, 2020

Introduction

10   Cooperative behavior can harm an individual's fitness and increase the fitness of its conspecifics or competitors (Axelrod and Hamilton, 1981). Nevertheless, cooperative behavior appears to oc-
12   cur in many non-human animals (Dugatkin, 1997), for example rats (Rice and Gainer, 1962) and birds (Krams et al., 2008). Evolution of cooperative behavior remains an important conundrum in
14   evolutionary biology. Three major theories have been proposed to explain the evolution of cooperative behavior:

16   1. *Kin selection* theory posits that natural selection can favor cooperation between related individuals. The importance of relatedness to the evolution of cooperation and altruism was shown by Hamilton
18   (1964). According to Hamilton, an allele that determines altruistic behavior should increase in frequency when the reproductive cost to the actor that cooperates, c, is less than the benefit to the
20   recipient, b, times the genetic relatedness between the recipient and the actor, r. This is also known as Hamilton's Rule:

$$c < b \cdot r. \tag{1}$$

22   There is an ongoing debate about the extent to which kin selection explains the evolution of cooperation
24   and altruism. For example, it has been suggested that it can explain the cooperative behavior of worker castes of eusocial insects like the honey bee. The most significant argument against kin selection is
26   that in some cases cooperation among unrelated individuals appears to have evolved (Wilson, 2005). According to Wilson (2005), this makes Hamilton's rule incomplete, although Foster et al. (2006)
28   reject this claim and argue that altruism cannot evolve without relatedness. They refer to Hamilton, who claimed that relatedness can arise without recent common ancestry. Wilson (2005) also criticizes
30   kin selection on the grounds that environmental or ecological factors are probably more important than relatedness in determining social actions. However, Foster et al. (2006) argue that kin selection
32   does not ignore ecology, and Hamilton's rule suggests that environmental factors causing a high benefit-to-cost ratio will favor cooperation.

34   2. *Reciprocity* entails that repeated interactions or individual recognition are key components of the evolution of cooperation. In *direct reciprocity* there are repeated encounters between the same two
36   individuals, and at every encounter each individual has a choice between cooperation and defection. Hence, it may eventually pay off to cooperate if it may cause your partner to cooperate in the future.
38   This game-theoretic framework, known as the *repeated prisoner's dilemma*, can only lead to the evolution of cooperation if the cost is less than the benefit b times the probability of another encounter
40   between the same two individuals, w,

$$c < b \cdot w. \tag{2}$$

42   Direct reciprocity assumes that both players are in a position to cooperate, but it can not explain cooperation in asymmetric interactions such as human philanthropy.
44   *Indirect reciprocity* has also been suggested to explain the evolution of cooperation. Nowak (2006) claims that direct reciprocity is like a barter economy based on the immediate exchange of goods,
46   while indirect reciprocity resembles the invention of currency. The currency that "fuels the engines" of indirect reciprocity is *reputation*. However, reciprocity assumes repeated interactions and therefore
48   has difficulty in explaining the evolution of cooperation if interactions are not repeated.

50   3. *Group selection* theory posits that cooperation is favored because it imparts an advantage to the whole group, if selection acts at the group level in addition to the individual level. A common model
  for group selection divides the population into groups in which there are cooperators that help other
52   group members and defectors that do not. Individuals reproduce proportionally to their fitness, and offspring are added to the same group as their parents. If a group reaches a certain size it can split
54   to two groups, so groups that grow faster will split more often. Groups with cooperators grow faster

than groups without cooperators, and cooperation can evolve in this model when the cost c is less than the benefit b times the ratio between the the number of groups m and the sum of m and the maximum group size n,

$$c < b \cdot \frac{m}{m+n}. \tag{3}$$

Group selection has been criticized by biologists who advocate a gene-centered view of evolution. It has also been criticized because for cooperation to take over the population it must have higher fitness than defection, while under group selection theory the fitness of cooperators at the individual level is lower than the fitness of defectors. Thus a trait with a lower fitness taking over the population is a contradiction Eldakar and Wilson (2011) reject this argument, claiming that it is a tautology and does not qualify as an argument against group selection. The distinction between individual and group selection requires a comparison of fitness differentials within and between groups in a multi-group population, and when a trait evolves by group selection, despite having lower fitness within a group, that group might have higher average fitness in competition with other groups, all things considered.

These theories assume that cooperation is genetically determined, which raises the question: *Is it possible that cooperation is determined by non-genetic factors?* Culture has significant impact on the behavior of humans (Ihara and Feldman, 2004; Jeong et al., 2018) as well as non-human animals (Bonner, 2018). Here we attempt to determine to what extent the evolution of cooperative behavior can be explained by *cultural transmission*, which allows an individual to acquire attitudes and behavioral traits from other individuals in its social group through imitation, learning, or other modes of communication (Cavalli-Sforza and Feldman, 1981; Richerson and Boyd, 2008). Feldman et al. (1985) introduced the first model for the evolution of altruism by cultural transmission. They showed that under vertical (parent-to-offspring) cultural transmission, Hamilton's rule does not govern the evolution of parent-to-offspring or sib-to-sib altruism.

Non-vertical transmission may be either horizontal or oblique: horizontal transmission occurs between individuals from the same generation, while oblique transmission occurs from adults to unrelated offspring. Evolution under either of these transmission models can be be more rapid than under pure vertical transmission (Cavalli-Sforza and Feldman, 1981; Ram et al., 2018). Lewin-Epstein et al. (2017) have demonstrated that non-vertical transmission, mediated by microbes that manipulate their host behavior, can help to explain the evolution of cooperative behavior. Interestingly, some of their analysis can be applied to cultural transmission, because models of cultural transmission are mathematically similar to those for transmission of infectious diseases (Cavalli-Sforza and Feldman, 1981).

Here we hypothesize that non-vertical cultural transmission can explain the evolution of cooperation. To test this hypothesis, we suggest a model in which behavioral changes are mediated by cultural transmission that can occur during social interactions. For example, if an individual interacts with a cooperative individual, it might learn that cooperation is a positive behavior and will be cooperative in the future. We develop cultural evolution models that include both vertical and non-vertical transmission of cooperation and investigate these models using mathematical analysis and simulations. Our results demonstrate cultural transmission can facilitate the evolution of cooperation even when genetic transmission cannot. These results suggest that further research on the evolution of cooperation should account for non-vertical transmission and that treatment of cooperation as a cultural, rather then genetic trait, can lead to a better understanding of this important and enigmatic phenomenon.

Models

We focus on the evolution of cooperation in a fully mixed population where cooperation is modeled using the *prisoner's dilemma*.

3

Consider a very large population whose members are characterized by their phenotype ϕ, which can be of two types, $\phi = A$ for cooperators or $\phi = B$ for defectors. An offspring inherits its phenotype from its parent via vertical transmission with probability v or from a random individual in the parental population via oblique transmission with probability $(1 - v)$. Following Ram et al. (2018), given that the parent phenotype is ϕ and assuming uni-parental inheritance, the conditional probability that the phenotype ϕ' of the offspring is A is

$$P(\phi' = A \mid \phi) = \begin{cases} v + (1 - v)p, & \text{if } \phi = A \\ (1 - v)p, & \text{if } \phi = B \end{cases}, \tag{4}$$

where $p = P(\phi = A)$ is the frequency of A among all adults in the parental generation.

Not all adults become parents due to natural selection, and we denote the frequency of phenotype A among parents with $\tilde{p}$. Therefore, the frequency $\hat{p}$ of phenotype A among juveniles (after selection and vertical and oblique transmission) is

$$\begin{aligned} \hat{p} &= \tilde{p}[v + (1 - v)p] + (1 - \tilde{p})[(1 - v)p] \\ &= v\tilde{p} + (1 - v)p. \end{aligned} \tag{5}$$

Individuals interact according to a prisoner's dilemma. Specifically, individuals interact in pairs; a cooperator suffers a fitness cost $0 < c < 1$, and its partner gains a fitness benefit b, where we assume $b > c > 0$. **Table 1** shows the payoff matrix, i.e. the fitness of an individual with phenotype ϕ_1 when interacting with a partner of phenotype ϕ_2.

	$\phi_2 = A$	$\phi_2 = B$
$\phi_1 = A$	$1 + b - c$	$1 - c$
$\phi_1 = B$	$1 + b$	1

Table 1: **Payoff matrix for prisoner's dilemma.** The fitness of phenotype ϕ_1 when interacting with phenotype ϕ_2. A is a cooperative phenotype, B is a defector phenotype, b is the benefit gained by an individual interacting with a cooperator, and c is the cost of cooperation. $b > c > 0$.

Social interactions occur randomly: two individuals with phenotype A interact with probability $\hat{p}^2$, two individuals with phenotype B interact with probability $(1 - \hat{p})^2$, and two individuals with different phenotypes interact with probability $2\hat{p}(1 - \hat{p})$.

Horizontal cultural transmission occurs between peers. It may occur between social partners with probability α, or between a random pair with probability $1 - \alpha$ (see **Figure 1**). Horizontal transmission is not always successful, as one peer may reject the other's phenotype. The probability for successful horizontal transmission of phenotypes A and B are T_A and T_B, respectively (**Table 2**).

Therefore, the frequency p' of phenotype A among adults in the next generation, after horizontal transmission, is

$$\begin{aligned} p' = {}& \hat{p}^2[\alpha + (1 - \alpha)(\hat{p} + (1 - \hat{p})(1 - T_B))] \\ & + \hat{p}(1 - \hat{p})[\alpha(1 - T_B) + (1 - \alpha)(\hat{p} + (1 - \hat{p})(1 - T_B))] \\ & + (1 - \hat{p})\hat{p}[\alpha T_A + (1 - \alpha)\hat{p}T_A] \\ & + (1 - \hat{p})^2[(1 - \alpha)\hat{p}T_A], \end{aligned} \tag{6}$$

which simplifies to

$$p' = \hat{p}^2(T_B - T_A) + \hat{p}(1 + T_A - T_B). \tag{7}$$

4

| Phenotype ϕ_1 | Phenotype ϕ_2 | Frequency | Fitness of ϕ_1 | $P(\phi_1 = A)$ via horizontal transmission: | |
				from partner, α	from population, $(1 - \alpha)$
A	A	$\hat{p}^2$	$1 + b - c$	1	$\hat{p} + (1 - \hat{p})(1 - T_B)$
A	B	$\hat{p}(1 - \hat{p})$	$1 - c$	$1 - T_B$	$\hat{p} + (1 - \hat{p})(1 - T_B)$
B	A	$\hat{p}(1 - \hat{p})$	$1 + b$	T_A	$\hat{p}T_A$
B	B	$(1 - \hat{p})^2$	1	0	$\hat{p}T_A$

Table 2: **Interaction frequency, fitness, and transmission probabilities.**

The frequency of A among parents (i.e. after selection) follows a similar dynamic, but also includes the effect of natural selection, and is therefore

$$
\begin{aligned}
\bar{w}\tilde{p}' = {} & \hat{p}^2(1 + b - c)\left[\alpha + (1 - \alpha)(\hat{p} + (1 - \hat{p})(1 - T_B))\right] \\
& + \hat{p}(1 - \hat{p})(1 - c)\left[\alpha(1 - T_B) + (1 - \alpha)(\hat{p} + (1 - \hat{p})(1 - T_B))\right] \\
& + (1 - \hat{p})\hat{p}(1 + b)\left[\alpha T_A + (1 - \alpha)\hat{p}T_A\right] \\
& + (1 - \hat{p})^2\left[(1 - \alpha)\hat{p}T_A\right],
\end{aligned}
\tag{8}
$$

where fitness values are taken from **Table 1** and **Table 2**, and the population mean fitness is

$$
\bar{w} = 1 + \hat{p}(b - c).
\tag{9}
$$

Equation 8 can be simplified to

$$
\begin{aligned}
\bar{w}\tilde{p}' = {} & \hat{p}^2(1 + b - c)\left(1 - (1 - \hat{p})(1 - \alpha)T_B\right) \\
& + \hat{p}(1 - \hat{p})(1 - c)\left(\hat{p}(1 - \alpha)T_B + 1 - T_B\right) \\
& + (1 - \hat{p})\hat{p}(1 + b)\left(\hat{p}(1 - \alpha) + \alpha\right)T_A \\
& + (1 - \hat{p})^2\hat{p}(1 - \alpha)T_A.
\end{aligned}
\tag{10}
$$

Results

Without Oblique Transmission

With only vertical and horizontal transmission, i.e. $v = 1$, Equation 5 becomes $\hat{p} = \tilde{p}$, and Equation 10 for the change in frequency p' of phenotype A among parents can be written as

$$
\begin{aligned}
\bar{w}\tilde{p}' = {} & \tilde{p}^2(1 + b - c)\left[1 - (1 - \tilde{p})(1 - \alpha)T_B\right] \\
& + \tilde{p}(1 - \tilde{p})(1 - c)\left[\tilde{p}(1 - \alpha)T_B + 1 - T_B\right] \\
& + \tilde{p}(1 - \tilde{p})(1 + b)\left[\tilde{p}(1 - \alpha) + \alpha\right]T_A \\
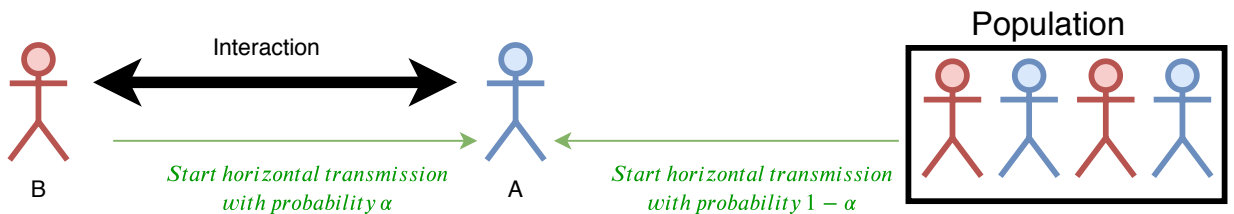& + (1 - \tilde{p})^2\tilde{p}(1 - \alpha)T_A.
\end{aligned}
\tag{11}
$$

Figure 1: **Cultural horizontal transmission.** Transmission occurs between interacting partners with probability α (left) or between two random peers with probability $1 - \alpha$.

We find the following result and corollaries.

Result 1: Vertical and horizontal transmission of cooperation. From Equation 11, if

$$c \cdot (1 - T_B) < b \cdot \alpha T_A + (T_A - T_B)[1 + b\tilde{p}(1 - \alpha)], \tag{12}$$

then $\tilde{p}' > \tilde{p}$, and the frequency of the cooperator phenotype A among parents increases every generation.

Corollary 1.1: Symmetric horizontal transmission. If $T = T_A = T_B$, then

$$c < b \cdot \frac{\alpha T}{1 - T}, \tag{13}$$

which can be seen as a cultural version of *Hamilton's rule* (Equation 1), where $r_H = \alpha T/(1 - T)$ can be thought of as *horizontal relatedness*. Therefore, if the cost c is less than the benefit b times the horizontal relatedness r_H, then cooperation will take over of the population (see **Figure 2a**).

Corollary 1.2: Complete correlation between transmission and cooperation.

In this case $\alpha = 1$, and horizontal transmission can only occur as a result of cooperative interactions. Therefore, Equation 12, which determines the conditions for evolution of cooperation, becomes

$$c \cdot (1 - T_B) < b \cdot T_A + (T_A - T_B). \tag{14}$$

This is equivalent to a result by Lewin-Epstein et al. (2017, eq. 1).

Equation 14 can be written as

$$1 - (1 - T_B)(1 - c) < T_A \cdot (1 + b), \tag{15}$$

which provides an interesting interpretation for the success of cooperation. Consider an interaction between two individuals: a cooperator and a defector. $(1 - T_B)(1 - c)$ is the probability that the cooperator remains cooperative and also reproduces. Therefore, $1 - (1 - T_B)(1 - c)$ is the probability that either the cooperator becomes a defector, *or* that it fails to reproduce. This is the effective cost for cooperation from this interaction. $T_A \cdot (1 + b)$ is the probability that the defector becomes cooperative and reproduces. This is the effective benefit for cooperation from this interaction. So, Equation 15 means that cooperation can evolve if the effective cost is less than the effective benefit.

Corollary 1.3: No correlation between transmission and cooperation.

In this case $\alpha = 0$, and horizontal transmission is entirely independent from cooperative interactions. Then, Equation 12 becomes

$$c \cdot (1 - T_B) < b \cdot \tilde{p} \, (T_A - T_B) + (T_A - T_B). \tag{16}$$

Therefore, since all the parameters are positive, cooperation cannot take over the population (and furthermore will become extinct) if cooperators do not have a horizontal transmission bias, i.e. if $T_A \leq T_B$. When such a bias does exist, $T_A > T_B$, then cooperation will evolve if $\tilde{p} > \tilde{p}^*$, where (see also **Figure 2b**)

$$\tilde{p}^* = \frac{c}{b} \cdot \frac{1 - T_B}{T_A - T_B} - \frac{1}{b}. \tag{17}$$

A sufficient condition for evolution of cooperation is that $\tilde{p}^* < 0$, namely

$$c < \frac{T_A - T_B}{1 - T_B}. \tag{18}$$

Without Vertical Transmission

With only oblique and horizontal transmission, i.e. $v = 0$, Equation 5 becomes $\hat{p} = p$ and Equation 7 becomes

$$p' = p^2(T_B - T_A) + p(1 + T_A - T_B), \tag{19}$$

which gives the following result.

Result 2: Oblique and horizontal transmission of cooperation. If there is a horizontal transmission bias in favor of cooperation, namely

$$T_A > T_B, \tag{20}$$

then $p' > p$, and the frequency of the cooperator phenotype among adults increases every generation. Therefore, cooperation will evolve if the cooperator phenotype has a horizontal transmission bias (see **Figure 2c**).

With Vertical and Oblique Transmission

In this case $0 < v < 1$, and the recursion system is more complex. Therefore, we focus on local stability, rather than global stability. To proceed, we note that Equation 5 can give $\hat{p}'$ as a function of both p' and $\tilde{p}'$, Equation 7 gives p' as a function of $\tilde{p}$, and Equation 10 gives $\tilde{p}'$ as a function of $\hat{p}$. Combining these equations, we find an equation for $\hat{p}'$ as a function of $\hat{p}$, see Appendix Appendix B. We then determine the equilibria, which are solutions of $\hat{p}' = \hat{p}$, and analyse their local stability.

We apply Equation 5, Equation 7, and Equation 10 to obtain the function $f(\hat{p})$, see Appendix Appendix B:

$$f(\hat{p}) = \bar{w}(\hat{p}' - \hat{p}) = \beta_1\hat{p}^3 + \beta_2\hat{p}^2 + \beta_3\hat{p}, \tag{21}$$

where

$$\begin{aligned}
\beta_1 &= \big[c(1-v) - b(1-\alpha v)\big](T_A - T_B), \\
\beta_2 &= -\beta_1 - \beta_3, \\
\beta_3 &= \alpha bvT_A - cv(1-T_B) + (T_A - T_B).
\end{aligned} \tag{22}$$

If $T = T_A = T_B$ then $\beta_1 = 0$ and $\beta_3 = -\beta_2 = \alpha bvT - cv(1-T)$. Therefore, $f(\hat{p})$ is a quadratic polynomial,

$$f(\hat{p}) = \hat{p}(1-\hat{p})\big[\alpha bvT - cv(1-T)\big]. \tag{23}$$

Clearly the only two equilibria are the fixations of either phenotype, $\hat{p} = 0$ and $\hat{p} = 1$. These equilibria are locally stable if $f'(\hat{p}) < 0$ (Appendix Appendix C). Therefore, we find the derivative,

$$f'(\hat{p}) = (1 - 2\hat{p})\big[\alpha bvT - cv(1-T)\big], \tag{24}$$

and investigate its sign at the equilibria,

$$\begin{aligned}
f'(0) &= \alpha bvT - cv(1-T), \\
f'(1) &= -\alpha bvT + cv(1-T).
\end{aligned} \tag{25}$$

Therefore with symmetric horizontal transmission, fixation of the cooperative phenotype ($\hat{p} = 1$) occurs under the same condition as Corollary 1.1, Equation 13.

In the general case where $T_A \neq T_B$, the coefficient β_1 is not necessarily zero, and $f(\hat{p})$ is a cubic polynomial. Therefore, three equilibria may exist, two of which are $\hat{p} = 0$ and $\hat{p} = 1$. By solving $f(\hat{p})/\big[\hat{p}(1-\hat{p})\big] = \beta_3 - \beta_1\hat{p} = 0$ we find the third equilibrium

$$\hat{p}^* = \frac{\beta_3}{\beta_1}. \tag{26}$$

7

212 Note that the sign of this cubic at positive (negative) infinity is equal (opposite) to the sign of β_1. If $T_A > T_B$, then

$$\beta_1 < [c(1 - \alpha v) - b(1 - \alpha v)](T_A - T_B) = (1 - \alpha v)(c - b)(T_A - T_B) < 0, \qquad (27)$$

216 since $c < b$ and $1 > \alpha v$, the sign of the cubic at positive and negative infinity is negative and positive, respectively. First, if $\beta_3 < \beta_1$ then $1 < \hat{p}^*$ and therefore $f'(0) < 0$ and $f'(1) > 0$, that is, fixation of the defector phenotype B is the only locally stable legitimate (i.e. between 0 and 1) equilibrium.
218 Second, if $\beta_1 < \beta_3 < 0$ then $0 < \hat{p}^* < 1$ and therefore $f'(0) < 0$ and $f'(1) < 0$, that is, both fixations are locally stable and $\hat{p}^*$ separates the domains of attraction. Third, if $0 < \beta_3$ then $\hat{p}^* < 0$ and therefore
220 $f'(0) > 0$ and $f'(1) < 0$, that is, fixation of the cooperator phenotype A is the only locally stable legitimate equilibrium.

222 Similarly, if $T_B > T_A$, then

$$\beta_1 > [c(1 - \alpha v) - b(1 - \alpha v)](T_A - T_B) = (1 - \alpha v)(c - b)(T_A - T_B) > 0, \qquad (28)$$

224 since $c < b$, and $1 > \alpha v$. So the sign of the cubic at positive and negative infinity is positive and negative, respectively. First, if $\beta_3 < 0$ then $\hat{p}^* < 0$ and therefore $f'(0) < 0$ and $f'(1) > 0$, that is,
226 fixation of the defector phenotype $A = B$ is the only locally stable legitimate equilibrium. Second, if $0 < \beta_3 < \beta_1$ then $0 < \hat{p}^* < 1$ and therefore $f'(0) > 0$ and $f'(1) > 0$, that is, both fixations are locally
228 unstable and $\hat{p}^*$ is a stable polymorphic equilibrium. Third, if $\beta_1 < \beta_3$ then $\hat{p}^* > 1$ and therefore $f'(0) > 0$ and $f'(1) < 0$, that is, fixation of the cooperator phenotype A is the only locally stable
230 legitimate equilibrium.

The following result summarizes these findings.

232 **Result 3: Oblique and vertical transmission of cooperation.** The cultural evolution of a cooperator phenotype will follow one of the following scenarios, depending on the horizontal transmission bias
234 $T_A - T_B$ and the coefficients β_1 and β_3:

1. *Fixation of the cooperative phenotype A,*

236   (a) if $T = T_A = T_B$ and $c < b \cdot \frac{\alpha T}{1-T}$, or

      (b) if $T_A > T_B$ and $0 < \beta_3$, or

238   (c) if $T_A < T_B$ and $\beta_1 < \beta_3$.

2. *Fixation of the defector phenotype B,*

240   (a) if $T = T_A = T_B$ and $c > b \cdot \frac{\alpha T}{1-T}$, or

      (b) if $T_A > T_B$ and $\beta_1 < \beta_3 < 0$, or

242   (c) if $T_A < T_B$ and $\beta_3 < 0$.

3. *Protected polymorphism, or co-existence of both phenotypes*, if $T_A < T_B$ and $0 < \beta_3 < \beta_1$.

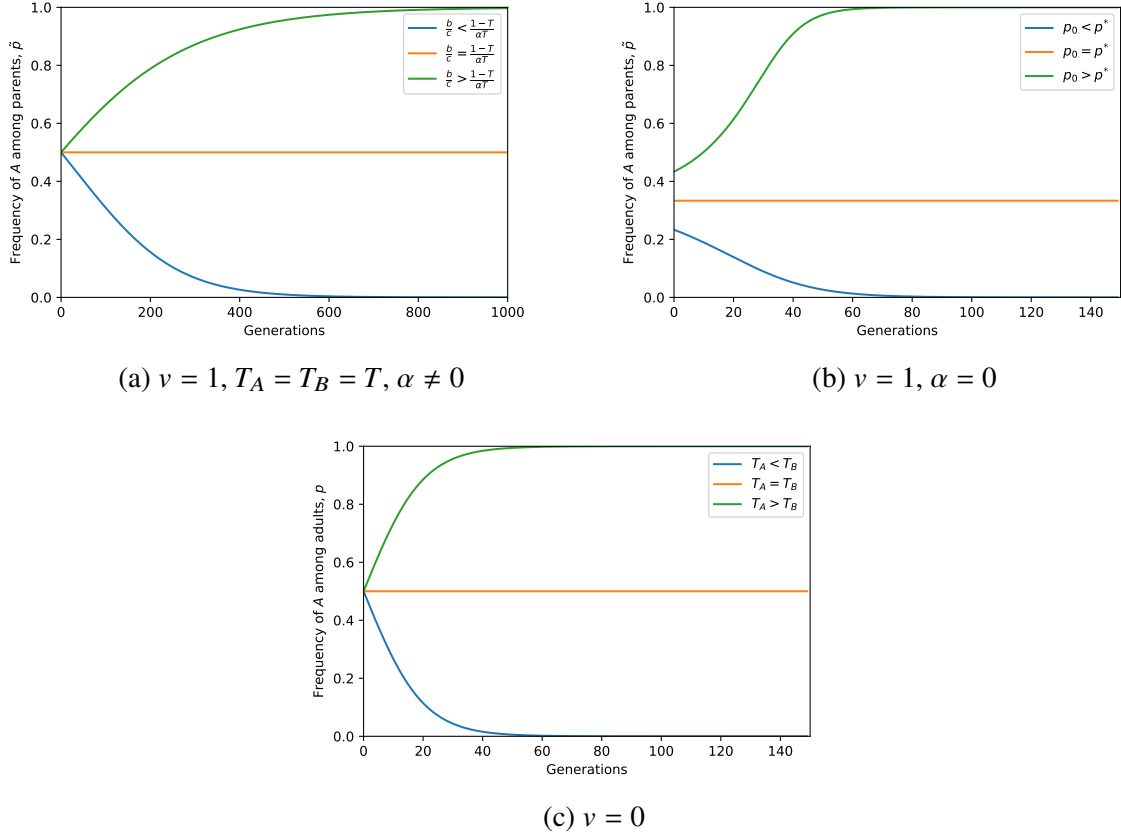244 4. *Fixation of either phenotype depending on initial frequency*, if $T_A > T_B$ and $\beta_3 < \beta_1$.

(a) $v = 1$, $T_A = T_B = T$, $\alpha \neq 0$

(b) $v = 1$, $\alpha = 0$

(c) $v = 0$

Figure 2: **Numerical results for cultural evolution of cooperation.** Shown are dynamics of **(a-b)** $\tilde{p}$, the frequency of parents with cooperative phenotype A; **(c)** p', the frequency of adults with cooperative phenotype A. The figure demonstrates fixation of cooperation (green), extinction of cooperation (blue)m and stable co-existence of cooperators and defectors (orange).

Discussion

We hypothesized that non-vertical transmission can explain the evolution of cooperation. We studied fully mixed and very large populations with a prisoner's dilemma payoff. We found that under horizontal and vertical cultural transmissions, if Equation 12 is satisfied, cooperation will take over fully mixed populations (Result 1). Under oblique and horizontal transmission, horizontal transmission bias for the cooperative phenotype is sufficient and necessary for evolution of cooperation (Result 2, Equation 20). Under a combination of vertical, oblique, and horizontal transmission the dynamics are further complicated. Yet, we find that cooperation can evolve and in some cases be maintained together with defection (Result 4). Importantly, our results demonstrate that cooperation can evolve even in a fully mixed population (i.e. in an unstructured population), without repeating interactions or individual recognition. These results significantly further our understating of the cultural evolution of cooperation.

This study was partially inspired by Lewin-Epstein et al. (2017). They hypothesised that microbes that manipulate their hosts to act altruistically can be favored by selection, and may play a role in the widespread occurrence of cooperative behavior. Indeed, it has been shown that microbes can mediate behavioral changes in their hosts (Dobson, 1988; Poulin, 2010). Therefore, natural selection on microbes may favor manipulation of the host so that it cooperates with others. Microbes can be transmitted *horizontally* from one host to another during host interactions, and following horizontal transfer, the recipient host may carry microbes that are closely related to the microbes of the donor host, even when the two hosts are (genetically) unrelated (Lewin-Epstein et al., 2017). Microbes can

9

also be transferred vertically, from parent to offspring, and a microbe that induces its host to cooperate with another host and thereby increases the latter's fitness will increase the vertical transmission of the microbes of the receiving individual. Kin selection among microbes could therefore favor microbes that induce cooperative behavior in their hosts, thereby increasing the transmission of their microbial kin.

Eshel and Cavalli-Sforza (1982) have studied a relevant model for the evolution of cooperative behavior under vertical transmission. Their model included *assortative meeting*, or non-random encounters. This assortment may be due, for example, to population structure or active partner choice. In this model, cooperative behavior can evolve if (Eshel and Cavalli-Sforza, 1982, eq. 3.2)

$$c < b \cdot m, \tag{29}$$

where b and c are the benefit and cost of cooperation, and m is the fraction of the population that meets an individual of the same phenotype rather than a random individual[1]. As a version of Hamilton's rule, m can be interpreted as the effective relatedness. In our Corollary 1.1 (Equation 13), cooperation evolves if $c < b \cdot \alpha T/(1 - T)$. So in our model $\alpha T/(1 - T)$ is the effective relatedness, which is affected by α, the correlation between transmission and interaction, and T, the horizontal transmission rate.

---

[1]In an extended model, which allows an individual to encounter N individuals before choosing a partner, the righthand side is multiplied by $E[N]$, the expected number of encounters (Eshel and Cavalli-Sforza, 1982, eq. 4.6).

Acknowledgements

References

Robert Axelrod and William D Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–
286   1396, 1981.

John Tyler Bonner. *The Evolution of Culture in Animals*. Princeton University Press, 2018.

288 Luigi Luca Cavalli-Sforza and Marcus W Feldman. *Cultural transmission and evolution: A quantitative approach*. Number 16. Princeton University Press, 1981.

290 Andrew P Dobson. The population biology of parasite-induced changes in host behavior. *The Quarterly Review of Biology*, 63(2):139–165, 1988.

292 Lee Alan Dugatkin. *Cooperation among Animals: An Evolutionary Perspective*. Oxford University Press on Demand, 1997.

294 Omar Tonsi Eldakar and David Sloan Wilson. Eight criticisms not to make about group selection. *Evolution*, 65(6):1523–1526, 2011.

296 Ilan Eshel and Luigi Luca Cavalli-Sforza. Assortment of encounters and evolution of cooperativeness. *Proc. Natl. Acad. Sci.*, 79(4):1331–1335, 1982. ISSN 2141-2502. doi: 10.5897/JPP2016.0416.
298   URL http://www.pnas.org/cgi/doi/10.1073/pnas.79.4.1331.

Marcus W Feldman, Luca L Cavalli-Sforza, and Joel R Peck. Gene-culture coevolution: models
300   for the evolution of altruism with cultural transmission. *Proceedings of the National Academy of Sciences*, 82(17):5814–5818, 1985.

302 Kevin R Foster, Tom Wenseleers, and Francis LW Ratnieks. Kin selection is the key to altruism. *Trends in Ecology & Evolution*, 21(2):57–60, 2006.

304 William D Hamilton. The genetical evolution of social behaviour. ii. *Journal of Theoretical Biology*, 7(1):17–52, 1964.

306 Yasuo Ihara and Marcus W Feldman. Cultural niche construction and the evolution of small family size. *Theoretical Population Biology*, 65(1):105–111, 2004.

308 Choongwon Jeong, Shevan Wilkin, Tsend Amgalantugs, Abigail S Bouwman, William Timothy Treal Taylor, Richard W Hagan, Sabri Bromage, Soninkhishig Tsolmon, Christian Trachsel, Jonas Gross-
310   mann, et al. Bronze age population dynamics and the rise of dairy pastoralism on the eastern eurasian steppe. *Proceedings of the National Academy of Sciences*, 115(48):E11248–E11255, 2018.

312 Indrikis Krams, Tatjana Krama, Kristine Igaune, and Raivo Mänd. Experimental evidence of reciprocal altruism in the pied flycatcher. *Behavioral Ecology and Sociobiology*, 62(4):599–605, 2008.

314 Ohad Lewin-Epstein, Ranit Aharonov, and Lilach Hadany. Microbes can help explain the evolution of host altruism. *Nature Communications*, 8:14040, 2017.

316 Aaron Meurer, Christopher P. Smith, Mateusz Paprocki, Ondřej Čertík, Sergey B. Kirpichev, Matthew Rocklin, AMiT Kumar, Sergiu Ivanov, Jason K. Moore, Sartaj Singh, Thilina Rathnayake, Sean
318   Vig, Brian E. Granger, Richard P. Muller, Francesco Bonazzi, Harsh Gupta, Shivam Vats, Fredrik Johansson, Fabian Pedregosa, Matthew J. Curry, Andy R. Terrel, Štěpán Roučka, Ashutosh Sa-
320   boo, Isuru Fernando, Sumith Kulal, Robert Cimrman, and Anthony Scopatz. Sympy: symbolic

computing in python. *PeerJ Computer Science*, 3:e103, January 2017. ISSN 2376-5992. doi: 10.7717/peerj-cs.103. URL https://doi.org/10.7717/peerj-cs.103.

Martin A Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563, 2006.

Robert Poulin. Parasite manipulation of host behavior: an update and frequently asked questions. In *Advances in the Study of Behavior*, volume 41, pages 151–186. Elsevier, 2010.

Yoav Ram, Uri Liberman, and Marcus W Feldman. Evolution of vertical and oblique transmission under fluctuating selection. *Proceedings of the National Academy of Sciences*, 115(6):E1174–E1183, 2018.

George E Rice and Priscilla Gainer. "Altruism" in the albino rat. *Journal of Comparative and Physiological Psychology*, 55(1):123, 1962.

Peter J Richerson and Robert Boyd. *Not by Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press, 2008.

Edward O Wilson. Kin selection as the key to altruism: its rise and fall. *Social Research*, pages 159–166, 2005.

Appendices

Appendix A

We start with Equation 11 and we want to investigate when $\tilde{p} < \tilde{p}'$, that is, when

$$\begin{aligned}
\bar{w}\tilde{p} < &\ \tilde{p}^2(1 + b - c)(1 - (1 - \tilde{p})(1 - \alpha)T_B) \\
&+ \tilde{p}(1 - \tilde{p})(1 - c)(\tilde{p}(1 - \alpha)T_B + 1 - T_B) \\
&+ \tilde{p}(1 - \tilde{p})(1 + b)(\tilde{p}(1 - \alpha) + \alpha)T_A \\
&+ (1 - \tilde{p})^2\tilde{p}(1 - \alpha)T_A.
\end{aligned} \tag{A1}$$

First divide by $\tilde{p}$ to obtain

$$\begin{aligned}
\bar{w} < &\ \tilde{p}(1 + b - c)(1 - (1 - \tilde{p})(1 - \alpha)T_B) \\
&+ (1 - \tilde{p})(1 - c)(\tilde{p}(1 - \alpha)T_B + 1 - T_B) \\
&+ (1 - \tilde{p})(1 + b)(\tilde{p}(1 - \alpha) + \alpha)T_A \\
&+ (1 - \tilde{p})^2(1 - \alpha)T_A.
\end{aligned} \tag{A2}$$

We know that the mean fitness $\bar{w} = 1 + \tilde{p}(b - c)$, thus Equation A2 becomes

$$\begin{aligned}
1 + \tilde{p}(b - c) < &\ \tilde{p}(1 + b - c)(1 - (1 - \tilde{p})(1 - \alpha)T_B) \\
&+ (1 - \tilde{p})(1 - c)(\tilde{p}(1 - \alpha)T_B + 1 - T_B) \\
&+ (1 - \tilde{p})(1 + b)(\tilde{p}(1 - \alpha) + \alpha)T_A \\
&+ (1 - \tilde{p})^2(1 - \alpha)T_A.
\end{aligned} \tag{A3}$$

which we simplify to inequality Equation 12.

Appendix B

We want to find the frequency of juveniles with phenotype A in next generation $\hat{p}'$ as a function of frequency of juveniles with phenotype A in the current generation $\hat{p}$. We start from eq. 5,

$$\hat{p}' = v\tilde{p}' + (1-v)p' \tag{B1}$$

Substituting p' using eq. 7 and $\tilde{p}'$ using eq. 10, we have

$$
\begin{aligned}
\hat{p}' =& \frac{v}{\bar{w}} \left\{ \hat{p}^2(1+b-c)\left[1-(1-\hat{p})(1-\alpha)T_B)\right]\right\} \\
& + \frac{v}{\bar{w}} \left\{\hat{p}(1-\hat{p})(1-c)\left[\hat{p}(1-\alpha)T_B + 1 - T_B\right]\right\} \\
& + \frac{v}{\bar{w}} \left\{\hat{p}(1-\hat{p})(1+b)\left[\hat{p}(1-\alpha)+\alpha\right]T_A\right\} \\
& + \frac{v}{\bar{w}}(1-\hat{p})^2\hat{p}(1-\alpha)T_A \\
& + (1-v)\hat{p}^2(T_B - T_A) + (1-v)\hat{p}(1+T_A - T_B),
\end{aligned}
\tag{B2}
$$

where $\bar{w} = 1 + \hat{p}(b-c)$.

We define $f(\hat{p})$ to be

$$f(\hat{p}) = \bar{w}(\hat{p}' - \hat{p}) \tag{B3}$$

Using *SymPy* (Meurer et al., 2017), a Python library for symbolic mathematics, we simplify Equation B3 to Equation 21 and Equation 22.

Appendix C

We show that $f'(\hat{p}^*) < 0$ is a sufficient condition for local stability an equilibrium $\hat{p}^*$. We will write $f(\hat{p})$ as a Taylor approximation around the equilibrium $\hat{p}^*$.

$$f(\hat{p}) = \bar{w}(\hat{p}' - \hat{p}) = f(\hat{p}^*) + f'(\hat{p}^*)(\hat{p} - \hat{p}^*) + O((\hat{p} - \hat{p}^*)^2) \tag{C1}$$

Where $O((\hat{p} - \hat{p}^*)^2)$ is big O notation. Denote, $\epsilon = \hat{p} - \hat{p}^*$ and $\epsilon' = \hat{p}' - \hat{p}$.

$$f(\hat{p}) = \bar{w}\epsilon' = f(\hat{p}^*) + f'(\hat{p}^*)\epsilon + O(\epsilon^2) \tag{C2}$$

Since $\hat{p}^*$ is an equilibrium $f(\hat{p}^*) = 0$ we get

$$f(\hat{p}) = \bar{w}\epsilon' = f'(\hat{p}^*)\epsilon + O(\epsilon^2) \tag{C3}$$

By definition of big-O, if $f'(\hat{p}*)$ is negative, then there exists some ϵ (i.e. some neighbourhood of $\hat{p}^*$) for which $f'(\hat{p}^*) + O(\epsilon)$ is also negative. Therefore,

$$\bar{w}\epsilon' = f'(\hat{p}^*)\epsilon + O(\epsilon^2) < 0 \tag{C4}$$

In addition, we know that $f(\hat{p}^*) = 0$ therefore, there exist some ϵ for which $f(\hat{p}) > -\delta$ for some $\delta > 0$. We choose delta to be $\delta = \frac{1}{\bar{w}}$, therefore $f(\hat{p}) = \bar{w}\epsilon' > -\frac{1}{\bar{w}}$. Therefore,

$$-\frac{1}{\bar{w}} < \bar{w}\epsilon' = f'(\hat{p}^*)\epsilon + O(\epsilon^2) < 0 \tag{C5}$$

Let's start simplify

$$-\frac{1}{\bar{w}} < \bar{w}\frac{\epsilon'}{\epsilon} < 0 \tag{C6}$$

Since $\bar{w} > 0$ we get that

$$-1 < \frac{\epsilon'}{\epsilon} = \frac{\hat{p}' - \hat{p}}{\hat{p} - \hat{p}^*} = \frac{\hat{p}' - \hat{p}^* + \hat{p}^* - \hat{p}}{\hat{p} - \hat{p}^*}$$

$$\frac{\hat{p}' - \hat{p}^*}{\hat{p} - \hat{p}^*} - 1 < 0$$

(C7)

From eq. C7 we get that

$$-1 < \frac{\hat{p}' - \hat{p}^*}{\hat{p} - \hat{p}^*} < 1$$

(C8)

We can apply absoulte value

$$|\hat{p}' - \hat{p}^*| < |\hat{p} - \hat{p}^*|$$

(C9)

Therefore, if $f'(\hat{p}^*) < 0$ there exist some neighbourhood of $\hat{p}^*$ where $\hat{p}'$ is closer to $\hat{p}$, which means that $f'(\hat{p}^*) < 0$ is sufficient condition for stability of $\hat{p}^*$.

Now let's assume by contradiction that $\hat{p}^*$ is also locally stable when $f'(\hat{p}^*) > 0$. When $f'(\hat{p}^*) > 0$ there exist some neighbourhood ϵ in which $f'(\hat{p}) > 0$ for every $\hat{p}$ in the neighbourhood.

If $\hat{p} > \hat{p}^*$ we get:

$$f(\hat{p}) = w\epsilon' > f(\hat{p}^*) = 0$$

(C10)

From eq. C10 we can see that $\hat{p}' > \hat{p} > \hat{p}^*$ and therefore, $\hat{p}^*$ is not stable.

If $\hat{p} < \hat{p}^*$ we get:

$$f(\hat{p}) = w\epsilon' < f(\hat{p}^*) = 0$$

(C11)

From eq. C10 we can see that $\hat{p}' < \hat{p} < \hat{p}^*$ and therefore, $\hat{p}^*$ is not stable.

We saw that if $f'(\hat{p}^*) < 0$ then $\hat{p}^*$ is stable, and if $f'(\hat{p}^*) > 0$ then $\hat{p}^*$ is unstable. Therefore, $f'(\hat{p}^*)$ is sufficient and necessary condition for equilibrium stability.