

# Cultural Transmission Can Explain the Evolution of Cooperative Behavior

Dor Cohen<sup>1</sup>, Ohad Lewin-Epstein<sup>2</sup>, Marcus W. Feldman<sup>3</sup>, and Yoav Ram<sup>a,\*</sup>

<sup>1</sup>School of Computer Sciences, IDC Herzliya, Herzliya, Israel

<sup>2</sup>School of Plant Sciences and Food Security, Tel Aviv University, Tel Aviv, Israel

<sup>3</sup>Department of Biology, Stanford University, Stanford, CA

\*Corresponding author: [yoav@yoavram.com](mailto:yoav@yoavram.com)

July 17, 2020

# Introduction

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 occur 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 evolutionary biology.

Cultural evolution is an evolutionary theory of individual and group-level change in beliefs, behaviors, and norms (Cavalli-Sforza and Feldman, 1981). 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). Cultural transmission 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). Here we attempt to determine to what extent cultural transmission can explain the evolution of cooperative behavior.

## Theories for evolution of cooperation

Three major theories have been proposed to explain the evolution of cooperative behavior.

**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 (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 recipient,  $b$ , times the genetic relatedness between the recipient and the actor,  $r$ . This is also known as Hamilton's Rule:

$$c < b \cdot r. \quad (1)$$

There is an ongoing debate about to what extent kin selection explains evolution of cooperation and altruism. It has been suggested that kin selection can explain the cooperative behavior of worker castes of eusocial insects like the honey bee. The most significant argument against kin selection is that in some cases cooperation among unrelated individuals appears to have evolved (Wilson, 2005). This makes Hamilton's rule incomplete according to Wilson (2005), although Foster et al. (2006) 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 kin selection on the grounds that environmental or ecological factors are probably more important than relatedness in determining social actions. Although Foster et al. (2006) argue that kin selection does not ignore ecology. Hamilton's rule suggests that environmental factors causing a high benefit-to-cost ratio will favor cooperation.

**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 individuals, where at every encounter, each player has a choice between cooperation and defection: if I cooperate now, you may cooperate later. Hence, it may pay off to cooperate. 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 between the same two individuals,  $w$ ,

$$c < b \cdot w. \quad (2)$$

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.

**Indirect reciprocity** has also been suggested as an explanation of this behavior. Nowak (2006) claims that direct reciprocity is like a barter economy based on the immediate exchange of goods, while indirect reciprocity resembles the invention of money. The money that “fuels the engines” of indirect reciprocity is reputation. However, reciprocity assumes repeated interactions and therefore has difficulty in explaining the evolution of cooperation if interactions are not repeated.

**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 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 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  $n$ ,

$$c < b \cdot \frac{m}{m + n}. \quad (3)$$

Group selection has been criticized by biologists who advocate a gene-centered view of evolution. Group selection has also been criticized because a trait such as cooperation evolves in the total population. For cooperation to take over the population it must have higher fitness, while under group selection the fitness of cooperators at the individual level is lower. 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.

All the above theories assume that cooperation is genetically determined, which raises the question, is it possible that cooperation is determined by non-genetic factors? 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. Recent work by Lewin-Epstein et al. (2017) also sheds some light on this question. They hypothesise 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 (Poulin, 2010; Dobson, 1988). 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 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.

## Cultural evolution of cooperation

Lewin-Epstein et al. (2017) have demonstrated that *non-vertical transmission* can help to explain the evolution of cooperative behavior. Non-vertical transmission may be either a horizontal or oblique: horizontal transmission occurs between individuals from the same generation, while oblique transmission occurs from an adult to an offspring. Evolution under either of these transmission models can be more rapid than under pure vertical transmission (Cavalli-Sforza and Feldman, 1981; Ram et al., 2018).

Here we suggest a model in which behavioral changes are mediated by cultural transmission that can occur during social interaction. For example, if an individual interacts with a cooperative individual, it might learn that cooperation is a positive behavior and will cooperate in the future. Some of the analysis made by Lewin-Epstein et al. (2017) 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 develop cultural evolution models that include both vertical and non-vertical transmission of cooperation and investigate these models using both mathematical analysis and simulations. We hypothesize that non-vertical cultural transmission can explain the evolution of cooperation.

## Models and Methods

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

Consider a very large population whose members are characterized by their phenotype  $\phi$ , which can be of two types,  $\phi = A$  for cooperators or  $\phi = B$  for defectors, with corresponding fitness values  $w_A$  and  $w_B$ , which depend on the frequency of the phenotypes (see below). 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  $\phi$  and assuming uni-parental inheritance, the conditional probability that the phenotype  $\phi'$  of the offspring is  $A$  is

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

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

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

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  $b > c$ . The following payoff matrix shows the fitness of an individual with phenotype  $\phi_1$  when interacting with a partner of phenotype  $\phi_2$  ( $b > c > 0$ ):

$$\begin{array}{cc} & \phi_2 = A & \phi_2 = B \\ \phi_1 = A & b - c & -c \\ \phi_1 = B & b & 0 \end{array}$$

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  $\alpha$ , 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 transmission of phenotypes  $A$  and  $B$  are  $T_A$  and  $T_B$ , respectively.

Table 1 contains the probability of first interactor to be  $A$  following interaction by interaction type. Therefore, the frequency  $p'$  of phenotype  $A$  among adults in the next generation, after horizontal transmission, is

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

which simplifies to

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

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

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

where fitness values are taken from Table 1, and

$$\bar{w} = \hat{p}w_A + (1 - \hat{p})w_B = 1 + \hat{p}(b - c) \quad (9)$$

is the population mean fitness. Eq 8 can be simplified to

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

Interaction Type	Frequency	Fitness	Probability of first interactor to be A following interaction	
			Horizontal transmission from partner, probability $\alpha$	Horizontal transmission from population, probability $1 - \alpha$
$AxA$	$\hat{p}^2$	$1 + b - c$	1	$\hat{p} + (1 - \hat{p})(1 - T_B)$
$AxB$	$\hat{p}(1 - \hat{p})$	$1 - c$	$1 - T_B$	$\hat{p} + (1 - \hat{p})(1 - T_B)$
$BxA$	$\hat{p}(1 - \hat{p})$	$1 + b$	$T_A$	$\hat{p}T_A$
$BxB$	$(1 - \hat{p})^2$	1	0	$\hat{p}T_A$

Table 1: Fitness and probability of the first interactor to be A following interaction by interaction type

## Results

We start by studying specific cases, for which we can derive general results, and then we use numerical simulation to analyze more complex cases.

### Without Oblique Transmission

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

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

We find the following result and corollaries.

**Result 1: Vertical and horizontal transmission of cooperation.** From eq. 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.

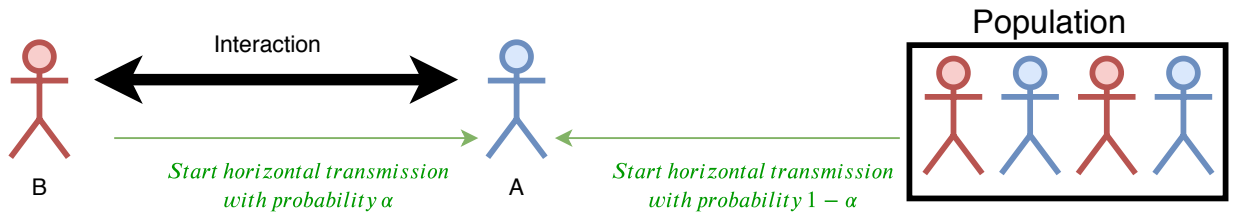


Figure 1: **Cultural horizontal transmission.** Transmission occurs between interacting partners with probability  $\alpha$  (left) or between two random peers with probability  $1 - \alpha$ .

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

$$c < b \cdot \frac{\alpha T}{1 - T}, \quad (13)$$

which can be seen as a cultural version of *Hamilton's rule* (eq. 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, eq. 12, which determines the conditions for evolution of cooperation, becomes

$$c \cdot (1 - T_B) < b \cdot T_A + (T_A - T_B). \quad (14)$$

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

Eq. 14 can be written as

$$1 - (1 - T_B)(1 - c) < T_A \cdot (1 + b), \quad (15)$$

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, eq. 15 means that cooperation can evolve if 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, eq. 12 becomes

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

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**)

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

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

$$c < \frac{T_A - T_B}{1 - T_B}. \quad (18)$$

## Without Vertical Transmission

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

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

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

$$T_A > T_B, \quad (20)$$

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, rather than global, stability. To proceed, we note that eq. 5 can give  $\hat{p}'$  as a function of both  $p'$  and  $\tilde{p}'$ , eq. 7 gives  $p'$  as a function of  $\hat{p}$ , and eq. 10 gives  $\tilde{p}'$  as a function of  $\hat{p}$ . Combining these equations, we find an equation for  $\hat{p}'$  as a function of  $\hat{p}$ . We then determine the equilibria, which are solutions of  $\hat{p}' = \hat{p}$ , and analyse their local stability: an equilibrium  $\hat{p}^*$  is locally stable when the derivative of  $f(\hat{p})$  at the equilibrium is negative,  $f'(\hat{p}^*) < 0$ .

We start with the simple case of symmetrical horizontal transmission,  $T = T_A = T_B$  and apply eqs. 5, 7, and 10 to obtain

$$f(\hat{p}) = \bar{w}(\hat{p}' - \hat{p}) = \hat{p}(1 - \hat{p})[\alpha bvT - cv(1 - T)]. \quad (21)$$

The equilibria are solutions of  $f(\hat{p}) = 0$ , or  $\hat{p}' = \hat{p}$ . It is easy to verify that fixation of either phenotype,  $\hat{p} = 0$  and  $\hat{p} = 1$ , is an equilibrium. Since the derivative of  $f'(\hat{p})$  is

$$f'(\hat{p}) = (1 - 2\hat{p})[\alpha bvT - cv(1 - T)], \quad (22)$$

then the condition for local stability of  $\hat{p} = 1$  is

$$f'(1) = -\alpha bvT + cv(1 - T) < 0, \quad (23)$$

which produces the following result.

**Result 3: Oblique and vertical transmission with symmetric horizontal transmission.** If horizontal transmission is symmetric,  $T = T_A = T_B$ , and if

$$c < b \cdot \frac{\alpha T}{1 - T}, \quad (24)$$

then fixation of the cooperator phenotype  $A$  is locally stable. The same condition was given in Corollary 1.1, Eq. 13.

We now turn to the general case where  $T_A \neq T_B$ . We have

$$f(\hat{p}) = \bar{w}(\hat{p}' - \hat{p}) = \beta_1 \hat{p}^3 + \beta_2 \hat{p}^2 + \beta_3 \hat{p}, \quad (25)$$

where

$$\begin{aligned} \beta_1 &= [c(1 - v) - b(1 - \alpha v)](T_A - T_B), \\ \beta_2 &= -\beta_1 - \beta_3, \\ \beta_3 &= \alpha bvT_A - cv(1 - T_B) + (T_A - T_B). \end{aligned} \quad (26)$$

Since  $f(\hat{p})$  is a cubic polynomial, three equilibria may exist, two of which are  $\hat{p} = 0$  and  $\hat{p} = 1$ . By solving  $f(\hat{p})/[\hat{p}(1 - \hat{p})] = \beta_3 - \beta_1 \hat{p} = 0$  we find the third equilibrium

$$\hat{p}^* = \frac{\beta_3}{\beta_1}. \quad (27)$$



Note that the sign of the cubic  $f(\hat{p})$  at positive (negative) infinity is equal (opposite) to the sign of  $\beta_1$ . If  $T_A > T_B$ , then

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

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. 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  $f'(0) > 0$  and  $f'(1) < 0$ , that is, fixation of the cooperator phenotype  $A$  is the only locally stable legitimate equilibrium.

Similarly, if  $T_B > T_A$ , then

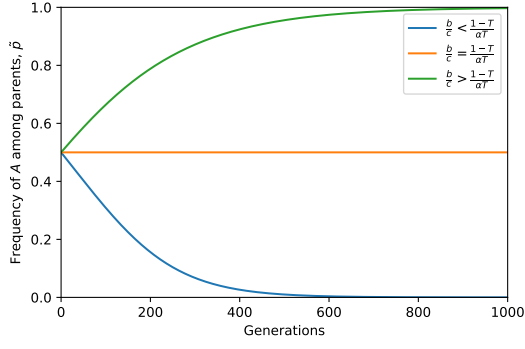
$$\beta_1 > [c(1 - \alpha v) - b(1 - \alpha v)](T_A - T_B) = (1 - \alpha v)(c - b)(T_A - T_B) > 0, \quad (29)$$

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, 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 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 legitimate equilibrium.

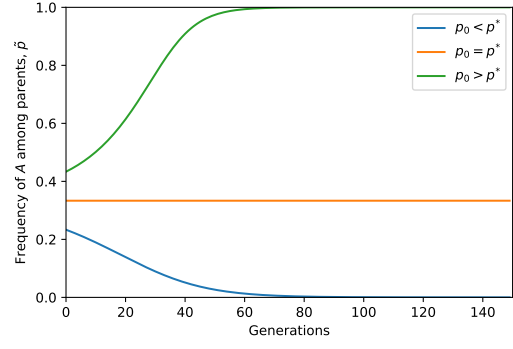
The following result summarizes these findings.

**Result 4: Oblique and vertical transmission of cooperation with asymmetric horizontal transmission.** The cultural evolution of a cooperator phenotype will follow one of the following scenarios, depending on the horizontal transmission bias  $T_A - T_B$  and the coefficients  $\beta_1$  and  $\beta_3$ :

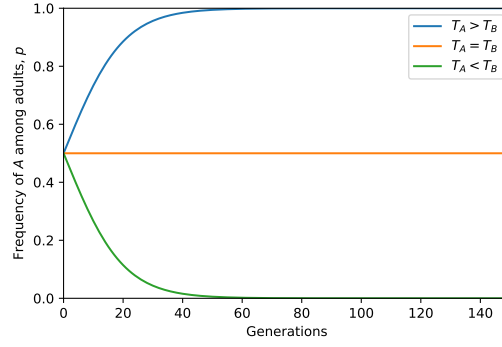
1. *Fixation of the cooperative phenotype A*, if  $T_A > T_B$  and  $0 < \beta_3$ , or  $T_A < T_B$  and  $\beta_1 < \beta_3$ .
2. *Fixation of the defector phenotype B*, if  $T_A > T_B$  and  $\beta_1 < \beta_3 < 0$ , or  $T_A < T_B$  and  $\beta_3 < 0$ .
3. *Protected polymorphism, or co-existence of both phenotypes*, if  $T_A < T_B$  and  $0 < \beta_3 < \beta_1$ .
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) and stable co-existence of cooperators and defectors (orange).

## Acknowledgements

We thank Lilach Hadany and Ayelet Shavit for discussions and comments. This work was supported in part by the Israel Science Foundation 552/19 (YR), and Minerva Center for Lab Evolution (YR).

## References

- Robert Axelrod and William D Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.
- John Tyler Bonner. *The Evolution of Culture in Animals*. Princeton University Press, 2018.
- Luigi Luca Cavalli-Sforza and Marcus W Feldman. *Cultural transmission and evolution: A quantitative approach*. Number 16. Princeton University Press, 1981.
- Andrew P Dobson. The population biology of parasite-induced changes in host behavior. *The Quarterly Review of Biology*, 63(2):139–165, 1988.
- Lee Alan Dugatkin. *Cooperation among Animals: An Evolutionary Perspective*. Oxford University Press on Demand, 1997.
- Omar Tonsi Eldakar and David Sloan Wilson. Eight criticisms not to make about group selection. *Evolution*, 65(6):1523–1526, 2011.
- Marcus W Feldman, Luca L Cavalli-Sforza, and Joel R Peck. Gene-culture coevolution: models for the evolution of altruism with cultural transmission. *Proceedings of the National Academy of Sciences*, 82(17):5814–5818, 1985.
- 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.
- William D Hamilton. The genetical evolution of social behaviour. ii. *Journal of Theoretical Biology*, 7(1):17–52, 1964.
- Yasuo Ihara and Marcus W Feldman. Cultural niche construction and the evolution of small family size. *Theoretical Population Biology*, 65(1):105–111, 2004.
- Choongwon Jeong, Shevan Wilkin, Tsend Amgalantugs, Abigail S Bouwman, William Timothy Treal Taylor, Richard W Hagan, Sabri Bromage, Soninkhishig Tsolmon, Christian Trachsel, Jonas Grossmann, 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.
- 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.
- Ohad Lewin-Epstein, Ranit Aharonov, and Lilach Hadany. Microbes can help explain the evolution of host altruism. *Nature Communications*, 8:14040, 2017.
- 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.

## Appendix A

In the section, we start with eq. 11 and we want to investigate when  $\tilde{p} < \tilde{p}'$ , that is, when

$$\begin{aligned}\bar{w}\tilde{p} &< \tilde{p}^2(1+b-c)(1-(1-\tilde{p})(1-\alpha)T_B) \\ &+ \tilde{p}(1-\tilde{p})(1-c)(\tilde{p}(1-\alpha)T_B + 1 - T_B) \\ &+ \tilde{p}(1-\tilde{p})(1+b)(\tilde{p}(1-\alpha) + \alpha)T_A \\ &+ (1-\tilde{p})^2\tilde{p}(1-\alpha)T_A\end{aligned}\tag{30}$$

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

$$\begin{aligned}\bar{w} &< \tilde{p}(1+b-c)(1-(1-\tilde{p})(1-\alpha)T_B) \\ &+ (1-\tilde{p})(1-c)(\tilde{p}(1-\alpha)T_B + 1 - T_B) \\ &+ (1-\tilde{p})(1+b)(\tilde{p}(1-\alpha) + \alpha)T_A \\ &+ (1-\tilde{p})^2(1-\alpha)T_A\end{aligned}\tag{31}$$

We know that the mean fitness  $\bar{w} = 1 + \tilde{p}(b-c)$  Thus eq. 31 becomes

$$\begin{aligned}1 + \tilde{p}(b-c) &< \tilde{p}(1+b-c)(1-(1-\tilde{p})(1-\alpha)T_B) \\ &+ (1-\tilde{p})(1-c)(\tilde{p}(1-\alpha)T_B + 1 - T_B) \\ &+ (1-\tilde{p})(1+b)(\tilde{p}(1-\alpha) + \alpha)T_A \\ &+ (1-\tilde{p})^2(1-\alpha)T_A\end{aligned}\tag{32}$$

Eq. 32 can be simplified to become inequality 12.