Reintroducing Kin Selection to the Human Behavioral Sciences

Hannah Rubin*†

Humans are often altruistic in a variety of contexts, even toward strangers they may never meet again. What explains this behavior? Many argue that kin selection cannot explain it but group selection can. Contra this common line of reasoning, I provide two ways that kin selection might help explain the evolution of broad-scope human altruism: in gene-culture coevolution and in a 'cultural' version of kin selection.

1. Introduction. Examples of altruistic behavior are everywhere we look. People donate blood and tip waiters in towns they will never return to. They systematically conserve fuel, recycle, and cut down on consumption to protect the environment—their contribution is individually costly in terms of time spent and material sacrifice but brings no real tangible benefit to them personally as one person alone cannot affect climate change. These altruistic actions are widespread and (to use Birch's [2017] term) 'broad-scope': they are often directed toward nonkin, members of large groups, and even strangers, with no expectation of reputational effects or repeated interactions.

In the time since Wilson and Sober (1994) discussed the (then) growing literature on group selection, reintroducing it as an important force in evolution, it has overshadowed the imagined importance of kin selection when it comes to explaining the emergence of broad-scope altruistic and cooperative behavior in human populations. One particularly influential line of reasoning in the human behavioral sciences maintains that kin selection, or interactions

Received March 2019; revised August 2019.

*To contact the author, please write to: Department of Philosophy, University of Notre Dame; e-mail: hannahmrubin@gmail.com.

†I would like to thank Simon Huttegger, Brian Skyrms, Cailin O'Connor, and audiences at Caltech and University of California, Irvine, for feedback on an early version of this article. Thanks also to Jonathan Birch for discussions on cultural kin selection. Finally, thanks to two anonymous reviewers for their helpful comments.

Philosophy of Science, 88 (January 2021) pp. 44–66. 0031-8248/2021/8801-0003$10.00
Copyright 2021 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

between close genetic relatives, cannot explain broad-scope human altruism. For example, Bowles and Gintis (2011) state that "because one of the distinctive aspects of human cooperation is that is extends far beyond the immediate family, we treat kin-based altruism only in passing" (49). Instead, this line of reasoning continues, group selection can explain broad-scope human altruism: in competition among groups, groups with more altruists will outcompete groups with fewer altruists.

I will argue that the quick dismissal of kin selection arguments overlooks important ways in which kin selection could have played an important role in shaping our altruistic behavior, even when it is directed toward nonrelatives. I begin with an explanation of why people reject kin selection, and the associated 'big mistake hypothesis', and why they accept group selection as an explanation of broad-scope altruism in humans (sec. 2). I then propose two ways in which kin selection could plausibly have been important in the evolutionary history of broad-scope human altruism, although in a different way than has been previously proposed. First, I provide a model using gene-culture coevolution to show how kin selection can play an important role, in combination with forces of cultural evolution, in filling a gap in explanations relying on group selection (sec. 3). Second, following a suggestion by Birch (2017), I show how a cultural, rather than genetic, form of kin selection could explain the evolution of altruism without an appeal to group selection (sec. 4). As will become clear, in arguing for the importance of kin selection, I will not argue against the use of group selection models in explaining human social behavior.

2. Prevailing Explanation of Human Altruism. What explains broad-scope human altruism, altruistic behavior directed at nonkin when there is no chance of the benefit being returned to the altruist (e.g., through reciprocal altruism or reputation gain)? As mentioned, this is the common conclusion: kin selection cannot explain broad-scope human altruism, but group selection can (Bowles and Gintis 2002; Boyd et al. 2003; Bowles and Gintis 2011; among others).[1] I discuss these two parts in turn, in sections 2.1 and 2.2.

Before this, though, a quick note about terminology. Some argue that kin selection is a special type of group selection (e.g., Sober and Wilson 1999), while others (e.g., Birch 2017, chap. 4) argue that they should be distinguished along the lines discussed here. Even if one believes that kin selection is a type of group selection, the two mechanisms described below remain distinct—and kin selection is dismissed as an explanation of broad-scope human altruism in favor of (a different type of) group selection.

1. See Burnham and Johnson (2005), who refer to those advancing this line of reasoning as 'the Collective', for further citations.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

2.1. Kin Selection. The arguments against kin selection explaining broad-scope human altruism are generally aimed at a particular type of argument invoking biological or genetic kin selection, which describes the evolution of social behaviors due to benefits of those behaviors falling on genetic relatives. Therefore, in this section, we will be discussing only genetic kin selection, leaving the discussion of the possibility of a cultural version of kin selection for section 4. We will focus on the evolution of *altruistic* behaviors, where an organism pays a cost c in order to bestow a benefit b on its social partner (which, in the context of kin selection, is a genetic relative who is likely to also have altruistic genes). It is important to distinguish altruism from *mutualism*, which is another form of cooperative behavior in which the social benefit of a trait does not come with a net cost. Different authors disagree about the relative importance of altruism versus mutualism in human evolutionary history, but it is generally accepted that broad-scope human altruism is something that is important to explain.

Although kin selection is a common explanation of altruism in nonhuman organisms, it is often dismissed fairly quickly when talking about altruism in humans. It is argued that kin selection can only explain altruism toward immediate family members, and since what we are interested in explaining is widespread altruism toward nonkin, kin selection cannot be explanatorily helpful (see, e.g., Bowles and Gintis 2011). Of course, humans also act altruistically toward kin, generally more so. There are some altruistic behaviors we only exhibit toward kin, and others are directed toward kin and nonkin alike. It is likely that the explanation for these kin-directed altruistic behaviors can proceed along lines different from the explanation for altruism toward nonkin and would not necessarily need to appeal to anything apart from kin selection.

This dismissal of kin selection for explaining broad-scope human altruism tends to be in response to a particular type of argument for the importance of kin selection, which is commonly referred to as the 'big mistake hypothesis' (e.g., Henrich and Henrich 2007; Bowles and Gintis 2011; Tomasello et al. 2012). This hypothesis asserts that altruism evolved in humans at a time when we lived in small kin groups in which, biologically, altruism was favored. In modern times, we are still altruistic because we have retained these genes for altruistic behavior even though they are no longer favored by evolution (see, e.g., Burnham and Johnson 2005). This argument has become popularly known as the big mistake hypothesis because it implies that all our altruistic actions toward nonkin are just big mistakes—they are just misfirings of our desire to help kin in a world where we no longer primarily interact with kin.[2]

2. The big mistake hypothesis actually includes misfires due to both desire to help kin and expectations of reciprocity, but we will focus on the kin selection part here.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

This argument is often dismissed quickly for a couple of reasons. First, other primates can distinguish kin from nonkin when deciding whether to behave altruistically, so it seems unreasonable that humans would not be able to do so. Second, it is argued that kin selection is unlikely to be important for explaining altruism in human societies because human groups were too large at the time we think modern human society started to evolve (around the late Pleistocene). That is, if humans were in groups of 50–100 people, they were not just interacting with close kin like parents and siblings. These large groups included many other individuals who were generally not highly related, so being altruistic toward people you interact with in general would not be favored under biological evolution (Fehr and Henrich 2003; Bowles and Gintis 2011, 94–95). In some modern hunter-gatherer societies, average relatedness is as low as .05 even in bands that number in the twenties (Hill et al. 2011).

2.2. Group Selection. Instead, group selection is proposed as the explanation of broad-scope altruism in modern humans. Here is an example of how a group selection argument might proceed. The basic idea is that groups whose members are altruistic will tend to outcompete other groups because they will more often survive things like environmental crises or attack by a predator. Even though altruism might be evolutionarily disadvantageous within a group, groups with more altruists do better, increase in size faster, and more often split into new groups that are full of altruists, so overall altruists increase in frequency. Figure 1 provides a simple illustration of this concept. The group of all altruists survives to reproduce two new groups of altruists, the group with no altruists dies out, and the group with a majority altruists survives to reproduce one group. Although the frequency of altruists decreases within this last group, the frequency of altruists increases overall.[3]

While group selection is somewhat contentious as a type of biological evolution, it is often argued that it is more reasonable as a form of cultural evolution in human groups. This is because group selection requires variation between groups (e.g., groups with various levels of altruists), and migration between groups tends to decrease this variation. While in nonhumans, there is generally little to maintain this variation, in humans, culture can maintain group differences. In human groups, there are norms for how to behave. Forces like insider bias, the tendency to interact with people within one's own group, means that people tend to learn behaviors from within their own group. Additionally, conformist bias reduces within-group differences: even if there is a fair bit of migration between groups (people leaving one group and joining

3. Group selection, in general, does not depend on group reproduction of any sort. What matters is that individuals with the altruistic trait increase in frequency, and their fitness is influenced by group membership (see, e.g., Okasha [2006] for a discussion).

https://doi.org/10.1086/710028 Published online by Cambridge University Press

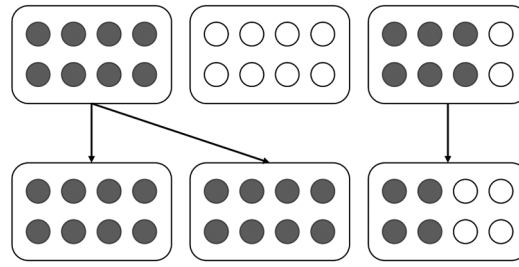


Figure 1. Example of group selection. Altruists are represented by filled circles, and nonaltruists, by open circles.

another), a group of altruists will remain a group of altruists because new members generally conform to the norms of the group.[4]

Authors arguing for group selection often add other parts to the story, such as punishment of those who do not adhere to norms or reproductive leveling that weakens within-group fitness differences, but the basic story is similar to that provided here. However, if we take this group selection argument seriously, and we think it provides a better explanation of altruism in humans than the big mistake hypothesis, there is still work to be done in explaining how the groups with varying levels of altruism arose in the first place.

One way to fill this gap in the argument is by appealing to kin selection, not in the same way as in the big mistake hypothesis but by showing that kin selection, as a type of biological evolution, can interact with forces of cultural evolution to produce groups with various levels of altruism. This will be discussed in section 3. There are, of course, other ways one might attempt to fill in this gap. Wilson and Dugatkin (1997), for instance, model altruism as a quantitative trait (on a scale from less to more altruistic). They show that when, during group formation, people shun others who are less altruistic than them, this can generate groups with various levels of altruism and can allow altruism to evolve. A variety of random processes can also lead to variation among groups (see, e.g., Boyd, Richerson, and Henrich [2011] and references therein).[5] The point here is not to argue that we must rely on kin selection but to show that, given what we can gather about the evolutionary history of human populations, kin selection provides a plausible mechanism by which we can find groups with various levels of altruism. Therefore, it should not be so quickly dismissed.

4. See, e.g., Okasha (2006, 159–60), Richerson and Boyd (2008, 162–63), and Bowles and Gintis (2011, 50–52) for a discussion.

5. See also Ellison (2000) and Kreindler and Young (2013) for a discussion on speed of convergence to different equilibriums in stochastic processes.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

Alternatively, one could avoid the appeal to group selection entirely and only appeal to kin selection, although it is a sort of kin selection different from that in the big mistake hypothesis. As will be discussed in section 4, even though genetic relatedness might have been low in human groups, cultural relatedness could still have been high (Birch 2017). That is, early humans would have tended to interact with others who had the same cultural traits—altruistic or not—even if they were not necessarily interacting with someone who had the same sort of genes. (When discussing cases in which altruism evolves because of high cultural relatedness, I will refer to this as cultural kin selection, while saving the term kin selection for interactions among genetic relatives.)

The models here may strike the reader as different from many existing models in the literature on human evolution. Many of these existing models are put forth in the service of providing a potential explanation of a particular observed human behavior, and, as such, they attempt to accommodate as many facts about human evolutionary history as possible, while still remaining both tractable and illustrative. The models here, by contrast, do not serve the purpose of defending or substantiating any particular story about how human altruism may have evolved. Rather, they identify relationships between different factors of interest (e.g., between relatedness, conformist bias, and levels of altruism). In doing so, they deliberately simplify away certain factors that would clearly affect the evolutionary process. This helps to isolate and identify factors that may be potentially relevant when attempting to put forth a hypothesis about what human evolutionary history may have looked like. For examples and defense of the usefulness of these sorts of models in economics, see Sugden (2000), and for recent use and defense of this sort of modeling practice in philosophy, see Rubin and O'Connor (2018).

To be clear, the models presented here will not license us to say anything like: "this is how altruism might have evolved in human populations." This is appropriate; the point here is not to put forth a competing explanation of human altruism. Rather, the point is to argue against the tendency to think of kin selection as unimportant and thus to not even consider it as a potential force at play when one puts forth an explanation of human altruism. That is, the argument in this article, which these models serve as evidence toward, is this: "kin selection should not be dismissed so quickly; it has a potential role to play in explaining human altruism, even toward nonrelatives."

3. Gene-Culture Coevolution. It is plausible that (genetic) kin selection was important in the evolutionary history of broad-scope human altruism. Here, I make a case for this claim, without arguing that all our altruistic actions are meant to be directed toward kin. I do this by providing a model in which kin selection interacts with the forces of cultural selection to provide groups with various levels of altruism—the starting point for models of group selection.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

While humans evolved biologically like any other species, cultural evolution has also been extremely important to the behaviors we see in modern human society. So, we have two types of evolution occurring at the same time: biological and cultural. *Gene-culture coevolution* provides a way to talk about how these two types of evolution occur and how they can interact with each other. In biological, or genetic, evolution, traits are influenced by genes that are passed on from parents to offspring. Traits that are beneficial will generally increase in frequency because people with those traits will tend to survive and reproduce more often.

By contrast, cultural evolution occurs when there is variation in behavior and some of these behaviors are adopted more often than others. In cultural evolution, traits are influenced by social learning and can be passed on in a variety of ways: from parents to their children, from any member of an older generation to a younger generation, between members of the same generation, and so on. This could occur because members of the group benefit from one behavior, and others observe the behavior and the resulting benefits and then imitate the behavior. But, in social learning, people do not always just pay attention to how beneficial a behavior is. It could be the case that individuals exhibit *prestige bias*, preferentially imitating group members who are most successful, whether or not the trait they are imitating is the source of that individual's success. Alternatively, it could be the case that one behavior is more prevalent than another, and people display some sort of *conformist bias*, meaning they adopt a behavior at least partially based on how common it is in the population. Conformist bias will be particularly important for the model presented here.

To demonstrate how gene-culture coevolution is important, here is one paradigmatic example of how it can occur. Before about 10,000 years ago, adult humans were not generally capable of processing the sugar in milk, called lactose. People could process the lactose as infants, but after they were weened off their mother's milk they stopped producing the necessary enzyme to metabolize it. After animals like cows started being domesticated, there was a steady source of milk readily available, and people who could process the nutrients in milk were favored by biological evolution. Those people had more nutrients in their diet, so they were able to survive and produce more offspring. This is well supported by historical and genetic evidence. For instance, there is a high frequency of people who can easily digest lactose in areas where dairying has been common for a long time and a low frequency in places where it has not historically been common (Feldman and Cavalli-Sforza 1989; Aoki 2001). This simple example shows how something that arose via cultural evolution affected our biological evolution, which is just one of the ways cultural and biological evolution can interact.

The various forms of social learning are also likely influenced to some degree by genetics that affect psychological predispositions. If groups are in a

https://doi.org/10.1086/710028 Published online by Cambridge University Press

relatively stable environment where, for example, a certain hunting technique remains successful generation after generation, then genes encoding for conformist bias are likely to evolve because it is much faster and less costly to simply do what everyone else is doing than to figure out the optimal hunting technique for yourself (Boyd and Richerson 1985, chap. 7; Henrich and Henrich 2007, chap. 2).

In section 3.1, I discuss a model of the evolution of altruism in which I take three claims about human evolutionary history to play an important role. It is important to emphasize, however, that these are far from hard-and-fast facts. Our evidence regarding the conditions of early human existence and evolution is often not clear or decisive, and we must make the best inferences we can using modern hunter-gatherers, archaeological data, mathematical models, behavioral experiments, and so on. Much of what is observed or claimed about human evolutionary history is quite tentative. So, we should think of the following claims (particularly the first and second) as claims that at least some people think there are good reasons to believe, and so these claims have a place in a model about a potential mechanism at play in human evolution:

1. Before the late Pleistocene, humans lived in smaller kin groups, which eventually grew in size to become the larger groups talked about in the group selection argument (Tomasello et al. 2012).
2. The benefit of conformist bias tends to increase as group size increases (Perreault, Moya, and Boyd 2012; Muthukrishna, Morgan, and Henrich 2016).[6] However, the level of conformist bias might be different in different groups at any given time for a variety of reasons. How fast groups adapt might vary, the behaviors influenced by conformist bias may be more or less crucial to their survival, or selective pressures on the groups might be stronger or weaker because they live in harsher versus more mild climates.
3. Relatedness tends to decrease as group size increases because there are more and more people in the group other than just immediate family members.[7]

These three claims about human evolutionary history are relevant to the evolution of altruism and will be taken into account in the model in the next section.

6. The idea is that, if five people exhibit a certain behavior, that is not as reliable as an indication of its quality as it would be if a group of 50 people exhibit the behavior, because it is more likely that the five people arrived at the behavior by chance.

7. Relatedness in groups depends on a variety of factors. See, e.g., Hamilton (1975) for an evaluation of how relatedness depends on both migration and group size.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

3.1. Model. Here is a basic description of how the evolution of genes and strategies occurs in this model. The terms *phenotype* or *strategy* refer to whether people choose to be an altruist, regardless of their genotype. Note that these altruistic behaviors are directed toward members of the group, regardless of whether they are kin. People start out with a strategy depending on their genotype; someone with altruist genes will likely be an altruist. People then interact within their group multiple times before reproducing. While they are interacting, people can update their strategies if they can see another strategy is doing better or (when there is conformist bias) if another strategy is more prevalent in the population. This is the process of cultural evolution. During these interactions they also accumulate material payoffs. The greater the material payoff people accumulate, the more offspring they have, and, thus, their genes (but not necessarily their strategy) will increase in the next generation. This is the process of biological evolution. We then start the process over with the next generation in which we have a period of cultural evolution and then one instance of biological reproduction.

I will explain in more detail the parts of the model in the following order: starting conditions, cultural evolution, biological evolution, and then group size increase over time and its effects on evolution. For ease of exposition, I assign particular values to many of the parameters, although none of these particular values will be crucial. Similar results can be obtained for a variety of parameter values. In the appendix (available online), many of the assumptions here are allowed to vary and qualitatively similar conclusions are drawn.

Initially, group size is small (at nine), to represent the situation in which people interact within a small kin group. Because altruism is biologically favored, a large proportion of group members have the altruistic genotype. A person with altruist genes is likely to be phenotypically, or culturally, altruist. For the results discussed here, the heritability of the trait is fairly high: a person with the altruistic gene is phenotypically altruist with probability .8, at least to start with. We have four 'phenogenotypes' to track in our evolutionary model, representing the combination of phenotype and genotype for each individual.

Within each biological generation, individuals interact a number of times and undergo cultural evolution. For each interaction, individuals will perform an altruistic action or not, depending on their phenotype. Then, after each interaction, the distribution of phenotypes (but not genotypes) evolves. This evolution occurs according to the discrete-time replicator dynamics, which captures the fact that if the value of altruism is greater than the value of nonaltruism at time t, then during the next time period, $t + 1$, altruism will increase in frequency:

$$x_a(t + 1) = \frac{x_a(t) \times v_a(x(t))}{\bar{v}(x(t))}, \tag{1}$$

https://doi.org/10.1086/710028 Published online by Cambridge University Press

where $v_a(t)$ represents the value of altruism at time t, and $\bar{v}(t)$ represents the average value of all traits in the population.

The value of altruism depends on $x_a(t)$, the frequency of altruists at time t, and the current level of conformist bias, $\mathcal{C}$. So, if $\mathcal{C}$ is high, the value of altruism will depend mostly on how frequent it is in the population, and if $\mathcal{C}$ is low, then the value of altruism will depend mostly on the material payoffs. If $u_a(x(t))$ is the material payoff gained in that interaction period based on the current distribution of strategies in the population and v_0 is the cultural analogue of 'background fitness' (see below) that accrues regardless of trait, then we can define the value of altruism similarly to Skyrms (2005):

$$v_a(x(t)) = v_o + \mathcal{C}x_a(t) + (1 - \mathcal{C})u_a(x(t)). \quad (2)$$

Remember that the level of conformist bias evolves over time, which I will return to shortly. (Note also that often conformist bias is taken to describe behavior in which a common trait is imitated more often than its frequency in the population, which this equation does not capture. See the appendix for a version of the model that captures this notion.)

We have thus far described how cultural evolution occurs within one biological generation. The model tracks the evolution of the population over 1,000 of these biological generations.[8] Throughout each generation, genes (or the humans who possess them) accumulate material payoffs based on their interactions. Then, one instance of biological reproduction occurs. This genetic evolution occurs according to an equation very similar to equation (1) but instead uses the fitness (the expected number of offspring, according to material payoffs) rather than the perceived value of a trait. If altruists have a greater fitness, the genes for altruism will increase in frequency.

We can calculate the fitness of altruists, f_a, using inclusive fitness. In calculating inclusive fitness, we look at the background fitness, f_0 (which accrues regardless of trait); the benefits, b, altruists confer on their relatives (weighted by how related they are, r); and the cost altruists have to pay, c. This gives us $f_a = f_0 + rb - c$. Nonaltruists do not pay any cost or confer any benefits, so their inclusive fitness is just $f_n = f_0$. In this model, we use information about how relatedness changes systematically over time, which makes inclusive fitness a useful framework for conceptualizing the evolutionary process (Rubin 2018).

It is now time to incorporate the three important claims about human evolutionary history discussed above. As noted, group size increases over time

8. In order to more accurately represent human populations, we also include some overlap between generations, $o = .45$. That is, not all the adults die when children are born; in any generation children make up a little over half the population.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

(claim 1). This is incorporated into the model as having group size, N, increase by 1 every 10 biological generations.[9]

As group size increases, the benefit of conformist bias increases (claim 2). As noted, how quickly the actual level of conformist bias will increase in the group depends on the strength of the selective pressures. This is incorporated into the model by tracking both an optimal and an actual level of conformist bias as groups increase in size. The optimal level of conformist bias is given by

$$C_{\text{optimal}} = \frac{N}{N + 100}. \tag{3}$$

This equation for the optimal level of conformist bias is chosen to represent a general trend; the specific form it takes is unimportant. Figuring out the actual equation for what the optimal conformist bias would be would require knowing a lot of particular details about the population, but here we are talking in very general terms about the evolution of traits influenced by conformist bias. The equation just needs to capture a couple of features. First, it needs to start out small to represent the observation that conformist bias is not as beneficial in smaller groups. Second, because of how it is incorporated into the equations describing cultural evolution, the level of conformist bias needs to stay between 0 and 1. These features are captured by equation (3), as shown in figure 2*a*.

The actual level of conformist bias depends on the selective pressures acting on the population, and, as we will see, the actual level of conformist bias is key to how much altruism can be sustained. We will talk about three cases: one in which actual conformist bias evolves so quickly that it is nearly optimal throughout the generations, another in which the conformist bias evolves very slowly, and an intermediate case, all shown in figure 2*b*.

As group size increases, relatedness also decreases (claim 3). We capture this by having relatedness change according to the following equation:

$$R = .5 - \frac{N - 9}{2(N - 9) + 50}. \tag{4}$$

Like the equation for C_{optimal}, knowing the exact equation for how relatedness evolves requires knowing specifics about reproduction, migration, and group structure, so this equation only captures some general features: relatedness starts out fairly high as most members of a small kin group are parents or siblings, and

9. Of course, there is no reason to assume that group size increases linearly rather than in some other particular way. However, it is important to note that, in this model, group size only affects the calculation of relatedness and level of conformist bias, as described below. The appendix looks at different assumptions regarding the form these two equations can take. If relatedness and conformism were to change in similar ways but for reasons different from group size increasing linearly, the conclusions based on the model would not change.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

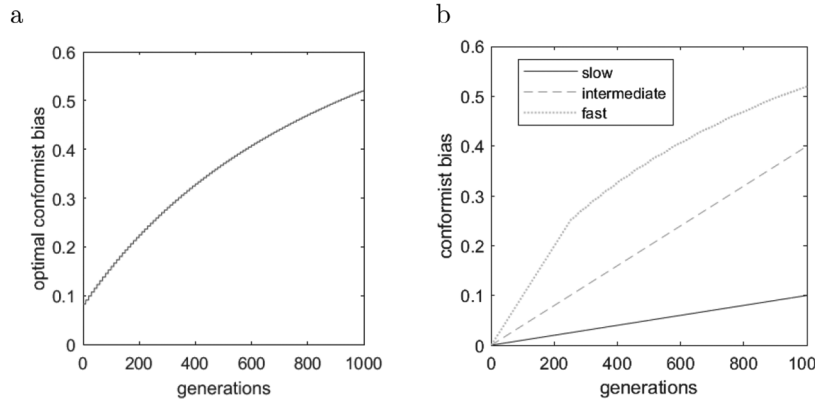


Figure 2. Optimal (*a*) and evolved (*b*) levels of conformist bias as group size increases. Color version available as an online enhancement.

it decreases over time to some level above 0, as shown in figure 3. Note that conformist bias and relatedness change are built in as assumptions of the model in such a way that they match the claims about the evolution of human groups described earlier. See the appendix for results with alternative equations for optimal conformist bias and relatedness that also have the features described here (but in which, e.g., conformist bias reaches its optimal level at lower group sizes and in which relatedness starts out lower and decreases faster).

3.2. Groups with Various Levels of Altruism. As mentioned, how much altruism can be sustained depends on the actual level of conformist bias in the population. Here, I present the results from the model described in the previous section in terms of the speed of evolution of conformist bias, corresponding to the three cases pictured in figure 2*b*. In the results presented here, $c = .1$, $b = .4$, and $v_0 = f_0 = 1$, but similar results can be obtained for a variety of values.

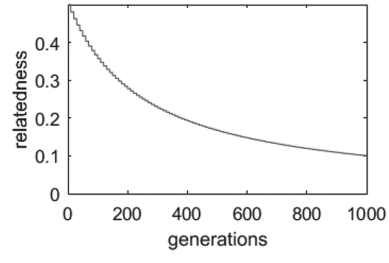


Figure 3. Relatedness as group size increases.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

Figures 4*a* and 4*b* show the evolution of the population when conformist bias increases quickly. We track how the frequencies of altruist genes and altruist phenotypes change over the generations in figure 4*a*. Initially, the altruistic gene is favored biologically because relatedness is high, but at around 250 generations, relatedness decreases to a point where these genes are no longer favored, and they slowly begin to decrease in frequency. For the altruistic phenotype, there is an initial decrease since conformist bias is low: people are born altruistic but fairly quickly learn that altruists have lower material payoffs than nonaltruists.

However, as conformist bias increases, the value of altruism increases, as shown in figure 4*b*. The perceived values of these behaviors is what affects cultural evolution. Initially, when conformist bias is low, the value of altruism is lower than the value of nonaltruism because it mostly depends on the material payoffs (and the material payoffs for altruists are lower than for nonaltruists). However, as conformist bias increases, people care more about how prevalent a trait is, and the value of altruism increases to the point where it is higher than the value of nonaltruism. (People are still born altruistic, so there is a high frequency of altruists.) Cultural altruism increases to a point where it can be maintained even when the genes for altruism begin to disappear (after relatedness has decreased to a point where altruism is no longer biologically favored). So, when conformist bias increases quickly, cultural altruism can be maintained for long periods of time.

What happens when conformist bias increases slowly? In this case, again, the altruistic gene is initially favored biologically because relatedness is high, but at around 250 generations, relatedness decreases to a point where these genes are no longer favored and begin to decrease in frequency. However, in this case, since conformist bias increases so slowly, the altruistic phenotype is not sustained culturally, as shown in figure 4*c*. While we start out with a high frequency of altruism, and some level of altruism is maintained for many generations (around 40%–50%), it eventually collapses as the genes for altruism disappear. This is because, as figure 4*d* shows, the value of altruism never increases to the point where it is greater than the value of nonaltruism. The value of altruism increases slightly as conformist bias increases, but once the altruistic genes disappear, the frequency of altruists drops and so does the value of altruism.

Finally, we can consider the case in which conformist bias increases at an intermediate speed. In this case, cultural altruism can be sustained for a short period of time. Again, the altruistic gene is initially frequent but drops out of the population as relatedness decreases. The altruistic phenotype increases in frequency for many generations, just as when conformist bias increased quickly, but it does not increase to levels as high. In this case, the value of altruism increases to where it is higher than the value of nonaltruism, but it does not increase quite as quickly as in the first case. This is shown in figure 4*f*. Because

https://doi.org/10.1086/710028 Published online by Cambridge University Press

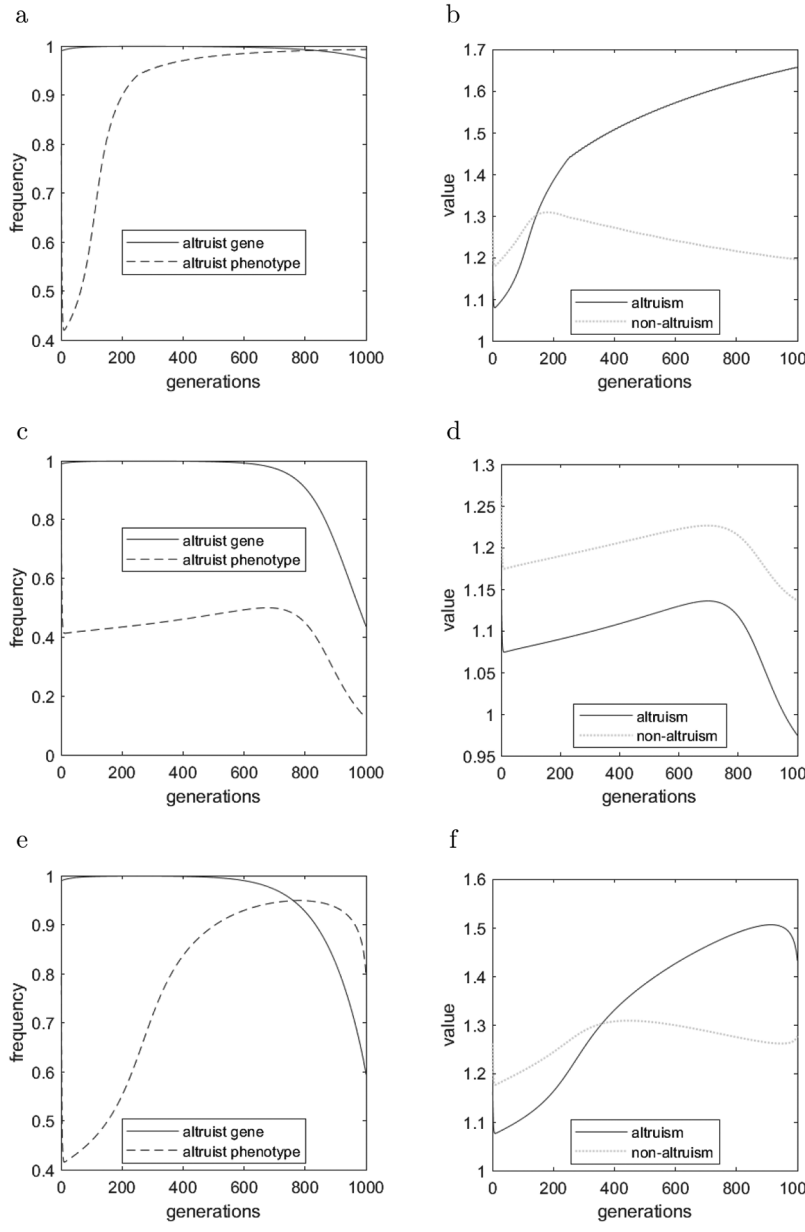


Figure 4. Evolution and value of strategies when conformist bias increases quickly (*a* and *b*), slowly (*c* and *d*), and at intermediate speed (*e* and *f*). Color version available as an online enhancement.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

phenotypic altruism never reaches a frequency as high as in the first case, it cannot be sustained for as long, and it eventually drops off toward the end of the 1,000 generations, as shown in figure 4*e*. Interestingly, even though the altruistic phenotype persists longer in the intermediate case than when conformist bias increases slowly, the gene for altruism disappears more quickly. This is because, with little conformist bias, the value of altruism is much lower, and genetic altruists switch to nonaltruism quickly enough during cultural evolution that there is little selection against them (they hardly ever pay the cost of altruism because they are nonaltruists for most of their lives).

So, when we vary how conformist bias evolves, this cultural evolutionary force can interact with kin selection in biological evolution in order to generate groups with different levels of altruism. These groups are exactly what is needed in order to start the group selection argument. There are a few important things to note. First, the altruism in these groups is not maintained indefinitely, so it would be necessary to combine this with group selection, or an explanation based on some other cultural force(s), in order to explain the broad-scope altruism we see today. (Although, see the appendix for a discussion of when altruism can be maintained indefinitely.) Second, the altruistic behavior in these models is always meant to be directed broadly toward everyone around the altruist, so the model provided here does not fall prey to the same objections leveled against the big mistake hypothesis. Finally, it does not rely on genetic relatedness remaining high as human groups increase in size to 50–100 people, the group size people generally agree the average human was in around the late Pleistocene.[10]

4. Cultural Kin Selection. Another approach, which does not rely on group selection at all, is to shift away from (genetic) kin selection and instead explain broad-scope human altruism as the result of cultural kin selection. While this general idea has been around for awhile (Cavalli-Sforza and Feldman 1981; Allison 1992; Lehmann and Feldman 2008; El Mouden et al. 2014), it is a relatively neglected approach (Birch 2017). The idea is that we can think of cultural traits as inherited (whether vertically, obliquely, or horizontally), and so in human groups, people can be highly culturally related, even when there is low genetic relatedness.

Birch (2017) argues for the potential importance of this selection process. In particular, he emphasizes the importance of horizontal transmission, where traits are transferred between individuals of the same generation, for human cultural evolution.[11] He argues that if horizontal transmission can generate

10. Specifically, there are 109 people at the end of the simulations above.

11. Interestingly, Birch is motivated to consider this because of the importance of horizontal transmission in explaining prosocial behavior in microbes.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

high enough cultural relatedness, altruism can evolve via cultural kin selection. The model provided here will show that, indeed, horizontal transmission of cultural traits can play this role.

Birch (2017) also argues for a diachronic conception of relatedness, where one takes into account the (genetic or cultural) similarity between individuals not just at the time of interaction but at other stages in their life cycle as well. One particularly interesting claim that follows from this is that high relatedness should promote the evolution of altruism, even when it is only generated after the relevant interactions; a person can increase the reproductive output of others, then make it so that those others spread that social behavior.[12] That the possibility of imitation after interaction is not something that is generally explored in models of cultural kin selection may have lead some to underestimate the potential importance of cultural kin selection.

4.1. Model. Here I provide a simple model to demonstrate the possibility that this sort of selection process can lead to the evolution of altruism. In this model, agents are arranged on a ring network where everyone has one neighbor (i.e., everyone interacts with those directly next to them). This network choice is not meant to be realistic—humans would certainly have interacted with more than two others, and there would be some form of clustering in the network. A ring was chosen because there is no nonarbitrary way to divide agents into groups, which shows that it is kin selection rather than group selection that explains the evolution of altruism in this model.[13] Note that since agents are arranged on a network and interact with their neighbors on the network, this model is not aimed at explaining altruism toward strangers. However, since the altruistic trait is assumed not to be linked to any particular genetics, it is aimed at explaining altruism toward non (genetic) kin.

We will track the evolution of cultural variants (as Birch [2017] refers to them), or *phenotypes* (as they were referred to in the previous section), which are not assumed to be connected to genotypes. Evolution in this model is very similar to the idealized life cycle described by Birch (204–5). The initial

12. Note that cultural selection in the context of this model is best thought of as still tied to biological reproduction in that cultural traits influence reproductive success. This is what Birch (2017, 197) calls type 1 cultural selection, or CS_1. Traits are passed on vertically from parents to offspring, although of course horizontal transmission also affects their evolution. This is opposed to CS_2, where traits affect cultural fitness (e.g., they influence the number of apprentices you have, who copy your behaviors, as opposed to your number of biological offspring).

13. Results similar to those described in sec. 4.2 have been obtained with a ring-lattice network (where agents are again arranged in a ring, but everyone interacts with those directly next to them and with the people one step away) and a square lattice (where agents are arranged on a torus and interact with their von Neumann neighbors).

https://doi.org/10.1086/710028 Published online by Cambridge University Press

distribution of strategies is random. In each round the agents are formed into a network, and the following stages occur:

1. *Horizontal Transfer 1* (HT1): Agents are chosen in a random order to observe the traits of their neighbors (the order in which they look at these neighbors is also random). Each time an agent observes a neighbor, it has a small probability of adopting or imitating the neighbor's trait, Pr_i. (It is possible that an agent changes a trait more than once during this stage.)
2. *Action*: Agents perform altruistic actions or do not, according to their trait. If the agent is an altruist, it acts altruistically toward each neighbor, paying a cost, c, and conferring a benefit, b, for each altruistic action.[14] Fitness for each agent is calculated by summing up costs paid and benefits accrued during this stage.
3. *Horizontal Transfer 2* (HT2): The same process as HT1 is repeated.
4. *Reproduction*: Reproduction occurs according to a Wright-Fisher model with selection. In this model, population size is constant. Agents are sampled from the population to reproduce, with replacement. The probability each agent is sampled is proportional to its fitness.[15] Offspring have the same trait as their parent at the time of reproduction (i.e., vertical transmission is perfect). So, even if an agent is not an altruist during the action stage, it will produce altruist offspring if they have gained the trait through the second round of horizontal transmission.

At the start of the next round, all the previous agents die, and a new network is formed with their offspring. This assumption is not meant to be a realistic description of evolution in human populations where there is overlap between generations and people would not randomly take a new place in the network, independent of their parents' ties. Instead, this is meant to isolate the effects of horizontal transfer and therefore demonstrate that the process of cultural kin selection can occur.[16]

14. This could be equally well described as paying a cost to produce a public good since all organisms have the same number of neighbors.

15. Each type has a background fitness of 1, to avoid negative payoffs.

16. It is worth flagging that the model in this section and that in sec. 3 serve slightly different purposes. While they are both simplified models, the model in sec. 3 was designed to match many more observations about human history. It showed how, out of a group of empirical claims relevant to the phenomenon, an effect emerges. This is because it was meant to show how two forces (including kin selection) could, in conjunction with each other, produce an effect (groups with various levels of altruism). The model here, by contrast, is meant to show how one force (cultural kin selection) can, by itself, lead to the evolution of broad-scope altruism. Therefore, it is designed to strip away other factors in order to isolate this one particular force.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

4.2. Results. The results shown here are for networks of 100 agents where evolution occurs over 300 generations.[17] The cost of the altruistic act was set to $c = .1$, and the benefit was varied from $b = 0$ to .6. The probability an agent would adopt their neighbor's trait was also varied, from $\Pr_i = 0$ to .2. Simulations were run 100 times to estimate the probability that a population would evolve to become composed entirely of altruists.

Figure 5*a* shows, first and foremost, that it is possible for altruism to evolve due to cultural kin selection. Perhaps unsurprisingly, as b or $\Pr_i$ increase, the evolution of altruism is more likely. It is important to note that in a Wright-Fisher model with selection, a certain amount of randomness influences the results. If there were no selection occurring, we would expect that 50% of the time, altruism would evolve. When altruism evolves more often, we can infer that, in general, selection tends to favor altruism, and when altruism evolves less often it is generally disfavored.

The results additionally show that relatedness need not be high at the time the altruistic actions are performed. Figure 5*c* shows that altruism can spread in a population even when horizontal transfer occurs only after an altruistic action has been performed. Since the network was randomly formed at the start of the round, there is no expected correlation between traits, and relatedness is on average 0 during the action stage. Of course, the existence of horizontal transfer both before and after the action stage is more conducive to the evolution of altruism than just having HT2, as comparing figures 5*a* and 5*c* shows. Additionally, while HT1 is more effective than HT2 (fig. 5*b* vs. 5*c*), having two bouts of horizontal transfer can be much more effective than only HT1 (fig. 5*a* vs. 5*b*).

Another way to conceptualize evolution in this model is to look at when we can expect (a cultural version of) Hamilton's rule to be satisfied. This rule states that we should expect altruism to be favored by natural selection whenever the (cultural) relatedness-weighted benefit of the trait exceeds its cost, or $r_v b - c > 0$, where r_v stands for relatedness among cultural variants. As Birch (2017) explains, for altruism to be favored it must be the case that those who are altruists during the action stage are sufficiently likely to help people who end up being altruists during the reproduction stage. We can therefore define cultural relatedness between an agent i and its social partner j as

$$c_v = \Pr(A_j^r | A_i^a) - \Pr(A_j^r | N_i^a),$$

where $\Pr(A_j^r | A_i^a)$ is the probability i's social partner is an altruist at the reproductive stage, given that i is an altruist at the action stage and $\Pr(A_j^r | N_i^a)$ is the

17. After 300 generations, 99.98% of simulations converged to the point where all of the population was one type or the other.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

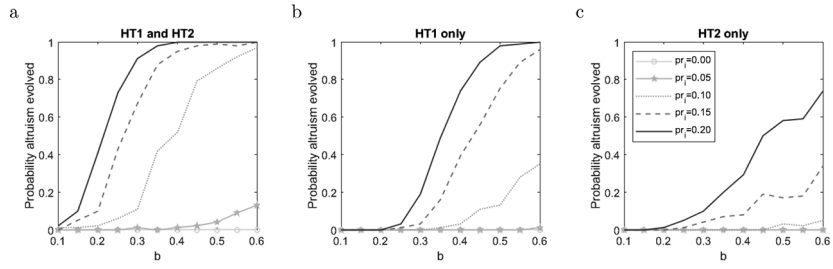


Figure 5. Proportion of simulations in which altruism evolved for different levels of b and $\Pr_i$, with (*a*) both stages of horizontal transmission included, (*b*) only horizontal transmission before the altruistic action is performed, and (*c*) only horizontal transmission after the altruistic action is performed. Color version available as an online enhancement.

probability i's social partner is an altruist at the reproductive stage, given that i is not an altruist at the action stage.[18]

Since the populations in this model are finite, the actual value of r_v can vary quite a bit. We can, however, get a sense of how often altruism will be favored by tracking how often r_v is high enough that $r_v b - c > 0$ is satisfied. Figure 6 shows the proportion of simulations in which altruism is favored.[19] For the values shown here, increasing the probability of imitation generally increases relatedness, and so altruism is more likely to be favored as this probability increases. As b increases, lower levels of r_v are needed to satisfy $r_v b - c$, and so, unsurprisingly, altruism is more likely to be favored as b increases. Comparing figures 6*a* and 6*c* confirms that altruism is also more likely to be favored with two stages of horizontal transmission versus including HT2 only. Finally, while HT1 makes it more likely altruism will be favored than HT2 (fig. 6*b* vs. 6*c*), having two bouts of horizontal transfer can make it much more likely than having HT1 only (fig. 6*a* vs. 6*b*).

The results just discussed are from a model in which people imitate and interact with only two neighbors. As mentioned, in reality, we interact with and imitate many more people than this. In fact, this may be part of the reason cultural kin selection is often overlooked in explaining human altruism—because cultural relatedness will tend to decrease as the number of people imitated increases, one might think that cultural kin selection will only be potentially important in special cases, for example, when there are a small number

18. Birch (2017) defines cultural relatedness similarly but in terms of covariances between traits since he derives it from the Price equation.

19. This was estimated by randomly forming 1,000 different networks with a random distribution of strategies, then measuring r_v on each and checking whether $r_v b - c > 0$.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

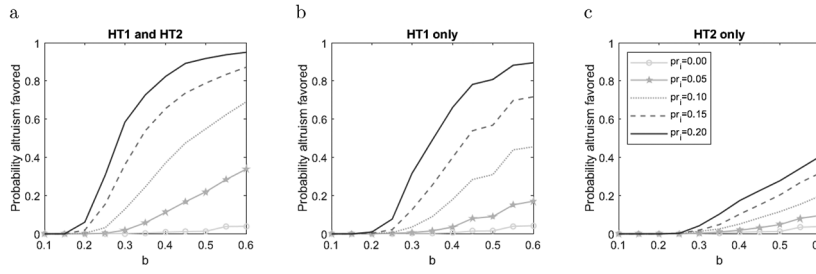


Figure 6. Proportion of simulations in which altruism is favored, according to $r_v b > c$ for different levels of b and $\Pr_i$, with (*a*) both stages of horizontal transmission included, (*b*) only horizontal transmission before the altruistic action is performed, and (*c*) only horizontal transmission after the altruistic action is performed. Color version available as an online enhancement.

of charismatic 'leaders' whom people imitate.[20] There are a couple ways to respond to this worry.

First, altruism can still be favored in this simple sort of model, even with a much larger number of neighbors. Figure 7 shows results for a ring lattice network where each agent is connected to the 10 closest nodes on each side, for a total of 20 neighbors. Figure 7*a* shows that, although it is much less likely, there is still a substantial probability that altruism is favored with this larger number of neighbors. For instance, when $b = .6$ and $\Pr_i$ is between .1 and .2, altruism is favored around a third of the time. Compare this to figure 7*b*, which only includes horizontal transmission before the trait and where altruism is much less likely to be favored. This shows that models that include only the possibility of imitation before interaction may severely underestimate potential for cultural kin selection to promote the evolution of altruism.

Second, one of the observations that motivated having two bouts of imitation was that the picture of a static level of relatedness generated before interaction was unrealistic. In fact, people have many instances of both interaction and imitation interspersed throughout their lifetime. The current model is still overly simplistic in this regard. One might imagine a more complicated model in which imitation of a large number of people would be intermixed with actually interacting with them. Even more realistic models would include other facts about human networks, including the facts that networks do not generally break apart and reform randomly, that there are often structural features like clustering, and so on. These more complicated models would not isolate cultural kin selection as the sole factor responsible for the evolution of

20. Thanks to an anonymous reviewer for pointing this out.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

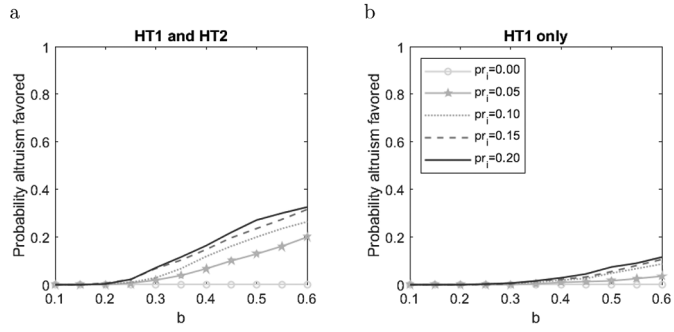


Figure 7. Proportion of simulations in which altruism is favored on a ring lattice with 20 neighbors, according to $r_v b > c$ for different levels of b and $\Pr_i$, when (*a*) both stages of horizontal transmission are included and (*b*) only horizontal transmission before the altruistic action is performed. Color version available as an online enhancement.

altruism.[21] However, such models could build on those presented here to include cultural kin selection as part of a potential explanation of human altruism.

The cultural kin selection explanation could, like the model in section 3, be compatible with the group selection hypothesis in answering why we have groups with various levels of altruism to begin with. It could be that the benefits of such behavior are higher in one group than another, based on some ecological conditions, leading to altruism evolving in one group but not another. It also could be that there is more horizontal transmission occurring in one group than other. And, it could just be that evolution is probabilistic, and while there is a mechanism for altruism to evolve, by chance sometimes a population ends up at the altruistic equilibrium and sometimes it does not.

5. Conclusion. We started out with this puzzle: Why do humans behave altruistically, and why do they direct their altruistic actions toward nonkin so frequently? Commonly, in resolutions of this puzzle, kin selection is dismissed in favor of explanations making use of group selection. I argued that kin selection should be reestablished as a potentially important explanatory factor in the human behavioral sciences by showing two ways in which it could be an important evolutionary force.

Since in humans, behavior is due to both our biological makeup and our culture, in the first model we looked to gene-culture coevolution to help us explain broad-scope human altruism. This model showed how kin selection

21. This is because, for example, reciprocity may come into play because being an altruist during one interaction period could make it more likely that one interacts with an altruist during a future interaction, or if there is sufficient clustering these clusters begin to look like groups.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

can interact with cultural evolution to give rise to groups of altruists. The second model leans more on the importance of cultural selection in providing a cultural version of the kin selection argument, where relatedness is now a measure of correlation between phenotypic traits rather than genotypes. These models can fill in the gaps in the group selection argument. More importantly, they show how kin selection can be an important force in explaining broad-scope human altruism, although not in a way that implies our altruistic actions are due to a misfiring of our desire to help kin.

This article makes a point often made in philosophy: just because one rejects the argument for a certain conclusion, one does not need to reject the conclusion. So while one might reject the current way some argue for the importance of kin selection, which relies on the big mistake hypothesis, we need not conclude that kin selection was not important to the evolution of broad-scope human altruism. Our evidence about the conditions of early human existence and evolution is often not clear or decisive, and we must make what inferences we can on the basis of modern hunter-gatherers, archaeological data, and so on. In the context of this uncertainty, it seems especially unwise to rule out an evolutionary force as potentially important without very strong reasons to do so. The models here show how kin selection is potentially important, by incorporating evidence about early human history or plausible assumptions about how humans interact with each other where appropriate. In comparing the merits of different explanations of broad-scope altruism in light of the data we do have, we should have all the (plausible) models on the table.

REFERENCES

Allison, P. D. 1992. "Cultural Relatedness under Oblique and Horizontal Transmission Rules." *Ethology and Sociobiology* 13 (3): 153–69.

Aoki, K. 2001. "Theoretical and Empirical Aspects of Gene-Culture Coevolution." *Theoretical Population Biology* 59 (4): 253–61.

Birch, J. 2017. *The Philosophy of Social Evolution*. Oxford: Oxford University Press.

Bowles, S., and H. Gintis. 2002. "Behavioural Science: *Homo reciprocans*." *Nature* 415 (6868): 125–28.

———. 2011. *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton, NJ: Princeton University Press.

Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson. 2003. "The Evolution of Altruistic Punishment." *Proceedings of the National Academy of Sciences* 100 (6): 3531–35.

Boyd, R., and P. J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Boyd, R., P. J. Richerson, and J. Henrich. 2011. "Rapid Cultural Adaptation Can Facilitate the Evolution of Large-Scale Cooperation." *Behavioral Ecology and Sociobiology* 65 (3): 431–44.

Burnham, T. C., and D. D. Johnson. 2005. "The Biological and Evolutionary Logic of Human Cooperation." *Analyse and Kritik* 27 (1): 113–35.

Cavalli-Sforza, L. L., and M. W. Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.

Ellison, G. 2000. "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution." *Review of Economic Studies* 67 (1): 17–45.

https://doi.org/10.1086/710028 Published online by Cambridge University Press

El Mouden, C., J.-B. André, O. Morin, and D. Nettle. 2014. "Cultural Transmission and the Evolution of Human Behaviour: A General Approach Based on the Price Equation." *Journal of Evolutionary Biology* 27 (2): 231–41.
Fehr, E., and J. Henrich. 2003. "Is Strong Reciprocity a Maladaptation? On the Evolutionary Foundations of Human Altruism." In *Mathematical Evolutionary Theory*, ed. P. Hammerstein, 55–82. Cambridge, MA: MIT Press.
Feldman, M. W., and L. L. Cavalli-Sforza. 1989. "On the Theory of Evolution under Genetic and Cultural Transmission with Application to the Lactose Absorption Problem." In *Mathematical Evolutionary Theory*, ed. M. W. Feldman, 145–73. Princeton, NJ: Princeton University Press.
Hamilton, W. D. 1975. "Innate Social Aptitudes of Man: An Approach from Evolutionary Genetics." In *Biosocial Anthropology*, ed. R. Fox, 133–53. London: Malaby.
Henrich, N., and J. P. Henrich. 2007. *Why Humans Cooperate: A Cultural and Evolutionary Explanation*. Oxford: Oxford University Press.
Hill, K. R., R. S. Walker, M. Božičević, J. Eder, T. Headland, B. Hewlett, A. M. Hurtado, F. Marlowe, P. Wiessner, and B. Wood. 2011. "Co-residence Patterns in Hunter-Gatherer Societies Show Unique Human Social Structure." *Science* 331 (6022): 1286–89.
Kreindler, G. E., and H. P. Young. 2013. "Fast Convergence in Evolutionary Equilibrium Selection." *Games and Economic Behavior* 80:39–67.
Lehmann, L., and M. W. Feldman. 2008. "The Co-evolution of Culturally Inherited Altruistic Helping and Cultural Transmission under Random Group Formation." *Theoretical Population Biology* 73 (4): 506–16.
Muthukrishna, M., T. J. Morgan, and J. Henrich. 2016. "The When and Who of Social Learning and Conformist Transmission." *Evolution and Human Behavior* 37 (1): 10–20.
Okasha, S. 2006. *Evolution and the Levels of Selection*. Oxford: Oxford University Press.
Perreault, C., C. Moya, and R. Boyd. 2012. "A Bayesian Approach to the Evolution of Social Learning." *Evolution and Human Behavior* 33 (5): 449–59.
Richerson, P. J., and R. Boyd. 2008. *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press.
Rubin, H. 2018. "The Debate over Inclusive Fitness as a Debate over Methodologies." *Philosophy of Science* 85 (1): 1–30.
Rubin, H., and C. O'Connor. 2018. "Discrimination and Collaboration in Science." *Philosophy of Science* 85 (3): 380–402.
Skyrms, B. 2005. "Dynamics of Conformist Bias." *Monist* 88 (2): 260–69.
Sober, E., and D. S. Wilson. 1999. *Unto Others: The Evolution and Psychology of Unselfish Behavior.* Cambridge, MA: Harvard University Press.
Sugden, R. 2000. "Credible Worlds: The Status of Theoretical Models in Economics." *Journal of Economic Methodology* 7 (1): 1–31.
Tomasello, M., A. P. Melis, C. Tennie, E. Wyman, and E. Herrmann. 2012. "Two Key Steps in the Evolution of Human Cooperation." *Current Anthropology* 53 (6): 673–92.
Wilson, D. S., and L. A. Dugatkin. 1997. "Group Selection and Assortative Interactions." *American Naturalist* 149 (2): 336–51.
Wilson, D. S., and E. Sober. 1994. "Reintroducing Group Selection to the Human Behavioral Sciences." *Behavioral and Brain Sciences* 17 (4): 585–608.

https://doi.org/10.1086/710028 Published online by Cambridge University Press