Explaining Human Social Aptitudes

Robert Boyd
Institute of Human Origins,
School of Human Evolution and Social Change
Arizona State University

Peter J. Richerson
Department of Environmental Science and Policy
University of California, Davis

Abstract

W.D. Hamilton in 1975 wrote a book chapter that constitutes his most extensive comments on human cooperation. In it he flagged the "tribal facies of social behavior" as the problem to be solved. He was well aware of the difficulty of extending his theory of inclusive fitness to the tribal scale. He mentions the idea that cultural processes might be responsible but expresses skepticism that culture could act against genetic fitness imperatives and sought genetic answers to the puzzle. We have explored the potential of culture to generate the stable variation necessary for selection at the level of tribes and other large human groups. We have modeled three forms of cultural group selection, and reviewed the ample empirical evidence that all three forms are important in humans. The reward and punishment systems in human societies can also create social selection on genes underlying human behavior. One of the critical factors in cultural evolution is that it can be faster than genetic evolution. Here we provide a simple model that illustrates why this is important to the evolution of the tribal facies.

Submitted for inclusion in *The Evolution of Cooperation and the Role of Hamilton's Rule*, edited by Ullica Segerstråle. June 2025.

Hamilton's approach to human cooperation

W.D. Hamilton's[1] fullest treatment of human cooperation was "The innate social aptitudes of man", a chapter in *Biosocial Anthropology*, edited by Robin Fox. It is a deep and interesting paper. The first part gives the first formal derivation of inclusive fitness from the Price equation, an approach that became one of the foundations of thinking about social evolution. This is followed by subtle, prescient thinking about the evolution of human sociality in which he considered both genetic and cultural explanations for the unusual scale of human cooperation.

Most evolutionary thinking about human cooperation has focused on small-scale cooperation in mobile foraging populations, especially food sharing. The large-scale cooperation in contemporary societies is often seen a resulting from "social aptitudes" that evolved in the context of such small-scale cooperation. With typical acuity, Hamilton zeroed in on what he called "the tribal facies of social behavior". Even hunting and gathering societies consist of ethnolinguistic units comprised of multiple fluid residential bands totaling, in aggregate, a few hundred to a few thousand people who sometimes cooperate on large scales. Hundreds of people engage in communal foraging, build shared facilities like drivelines, irrigation canals and fish traps, engage in large-scale habitat modification, and participate in warfare.[2] He noted that Darwin proposed a tribal scale group selection hypothesis to account for our unusual pattern of social behavior, and warns that we should be skeptical of such hypotheses because his and other's formal modeling work had shown that it is very hard to maintain genetic variation in between tribal scale groups. Large-scale cooperation is very rare among vertebrates, and not easily explained as altruism toward close kin or reciprocity.[3]

Hamilton offers several different kinds of explanations for tribal scale cooperation. He devotes much of the essay to developing an explanation based on inclusive fitness. He points out that developments in population genetics in the 1960's suggest that spatial structure can lead to substantial levels of relatedness and considers a mixed stepping stone—island model. Here relatedness within tribal scale groups, he argues, might approach 0.5 so that selection would favor lots of altruistic cooperation on tribal scales.

Problems with the inclusive fitness approach

Forty-five years later, this explanation seems less likely. Measurements of average relatedness within hunter-gatherer groups range from less than 0.01 to about 0.05 indicating the benefits would need to be somewhere between 20 and 100 times the cost.[4] It is possible to generate more altruism by adding other assumptions to the model[5] that amplify the effect of kin selection but these require some other mechanism to explain. Moreover, although Hamilton certainly was aware of the impact of local population regulation on the evolution of altruism, biologists have come to realize that local population regulation, an essential component of Hamilton's model tends to cancel the effects of increased relatedness. In the simplest model the two effects exactly balance and there is no selection for altruism.[6] Other models of population regulation[7] allow more altruism, but it is still limited compared to a model without population regulation. Finally, biologists have come to realize that weak selection is central to the application of Hamilton's rule. This means that it applies to relatedness due to recent common descent, but not to relatedness that builds up over the very long term in populations due to population structure.[8] For example, genetic F_{ST} for continental human populations averages around 0.14,[9] but this does not generate altruism toward other members of these populations because selection is weak.[10]

Tribal scale cooperation need not be based on altruism. In the last section of his paper, Hamilton focuses on “reciprocation and social enforcement”. He argues, correctly we think, that reciprocation can’t explain tribal scale cooperation, and instead argues that in n-person interaction direct sanctions better support cooperation because, as is expressed in this passage, third parties share an interest in their mutual benefit.

> …So the detection of cheating arouses indignation in everyone except the accused and everyone sees a benefit to both group and self in trying to punish the cheater…[11]

If the cost of being punished is greater than the cost of cooperating, then self-interested individuals will do what’s necessary to avoid punishment---invest in shared facilities, cooperate in large-scale communal hunting or participate in warfare. The problem is why should anybody punish? As Hamilton says, everyone will see that punishing cheaters is beneficial, but that does not mean that it will pay individuals to punish. The problem is that punishment itself is a public good, and in large groups the private benefit to the punisher of his or her own punishment is unlikely to pay the cost of administering it. Hamilton is a bit cryptic here, but seems to say that some genuine altruistic motivation could motivate the formation of coalitions that would provide the necessary punishment.

Another, more plausible, assumption is that people punish because punishment itself is normative.[12] People punish when it is expected because to fail to do so would lead to punishment or a loss of social standing. Such “normative enforcement” requires that norms specify when imposing sanctions is legitimate---how guilt is adjudicated, who is to impose the sanctions, what kinds of sanctions are appropriate. Otherwise sanctions are just another form of interpersonal conflict subject to retaliation, competing alliances and all the rest.[13] There is no limit to the size of cooperating groups with norms of this kind, in theory. However, empirical evidence indicates that normative enforcement without formal coercive institutions like courts and police can support collective action involving hundreds of unrelated individuals but not tens of thousands. Sarah Mathew[14] has studied cattle raiding by the Turkana, nomadic pastoralists who lack formal coercive institutions. Turkana raiding parties are large, averaging about 300 warriors. The individual costs are substantial. Free riders who desert, exhibit cowardice, or cheat on the division of the loot are punished by members of their own age set. A vignette study suggests that individuals who fail to punish will themselves suffer negative consequences, as will individuals who punish excessively or without sufficient evidence.[15]

Normative punishment can support any norm as long as the cost of being punished is greater than the benefit of violating the norm and the cost of enforcing the norm. Thus, a vast range of norms can be sustained as stable social equilibria. To explain altruism, we have to have a mechanism that causes mutually beneficial norms to become common. One such mechanism is competition between groups. Hamilton noted that Darwin proposed a tribal-scale group selection hypothesis to account for our unusual pattern of social behavior. He asserted that we should be skeptical of such hypotheses because modeling work had shown that it is difficult to maintain genetic variation among tribal scale groups.

> …cultures too have to survive and will hardly do so when they undermine the viability of their bearers. Thus, we should expect the genetic system to have various inbuilt safeguards and provide not a bank sheet for individual cultural development but a sheet at least lightly scrawled with certain tentative outlines.[16]

Here Hamilton's argument is exactly backwards. Of course, our psychology tends to cause us to adopt behaviors that are good for us. In a world without norms, self-interest undermines mutually beneficial cooperation and thus undermines the viability of societies. Normative punishment can finesse this problem. If cooperation is normative, then individuals will choose to behave cooperatively because now it is supported by our evolved selfish or nepotistic tendencies.

The problem is that normative punishment can maintain any behavior. Why should we believe that norms tend to support cooperative behavior? We have suggested that competition between groups with different norms for space, for economic success and for migrants will lead the spread of norms that enhance group competitive ability.[17] We have called this process cultural group selection, a decision that has led to much confusion because many people[18] think that group selection only refers to settings in which there is a conflict between the individual and the group. But this usage is perverse. Norms vary among groups. In a competitive world, some groups succeed and replace other groups, at least culturally. By any sensible standard this is group selection. (It is also etymologically valid---the term group selection began with Sewall Wright, not Wynne-Edwards, and he had in mind a process much like cultural group selection). Cultural group selection is about equilibrium selection. The assumption is that within each group it is in individual's self-interest to obey and enforce the local norm.

Three other kinds of explanations for punishment have been suggested. All of these require some mechanism like cultural group selection to reliably produce group beneficial norms. Punishing may signal an otherwise hard to observe individual characteristic that is desirable in social partners, so that those who impose sanctions are benefited,[19] but connection between the signal and group benefit is accidental. Participants may agree to allocate a big enough share of the benefits to leaders so that it is in the leader's interest to impose sanctions.[20] Something additional is necessary to motivate leaders to sanction in the interest of the group. Or, noncooperators can be deprived of the protections that are normally due to members of the society, who are then victimized by others for selfish reasons.[21] Here something must be added to explain why people who do not behave in the group interest are no longer protected from predation.

Hamilton expressed skepticism about the idea that cooperative behaviors could spread culturally. Surely genes would dominate any tendency for culture to violate considerations of genetic fitness. In 1975 the study of cultural evolution along Darwinian lines was in its infancy. There were a handful of programmatic essays[22] and Cavalli-Sforza and Feldman[23] had introduced the idea of modeling culture using tools borrowed from population genetics. In the ensuing years, the subject of cultural evolution has grown into a major enterprise and the possible role of culture playing a leading role in human cooperation is now well developed. Cultural inheritance has different properties than genetic inheritance and this means that the processes of cultural evolution do not precisely track those of genetic evolution.[24] Two related differences are particularly salient for our discussion here. First, cultural evolution is generally faster than genetic evolution[25] because learning and other creative processes mean that innovations, unlike genetic mutations, are not entirely random with respect to fitness and because social learners can selectively acquire adaptive variants they observe. Mutation and natural selection are typically weak forces in genetic evolution. Mutation is random with respect to fitness, and high rates of mutation impose a maladaptive load on populations. Natural selection operates by selective mortality and fecundity leading to substantial rates of mortality or sub-fecundity the demographic consequences of which can be serious in small low fecundity populations characteristic of most of human evolutionary history. At the same time, migration is normally a fairly strong force in genetic evolution. Hence, exceptional cases like Argentine ants aside,[26] genetic variation between neighboring groups of distantly related individuals is dominated by migration and remains small. In the case of cultural evolution, innovation can be better than random and choice-based selection allows individuals to acquire adaptive innovations without any direct demographic consequences. Also, favorable innovations can spread across social networks within a biological generation, circumventing the need for multiple generations to build up the frequency of favored cultural variants by natural selection. Therefore, contrary to Hamilton's intuition, group (or multi-level) selection acting on cultural variation is a plausible explanation for the tribal facies of human behavior.

Theoretical models

Theoretical work suggests that selection among groups is likely to be a more important process in cultural evolution than in genetic evolution. The argument that human cooperation between non-relatives is supported by selection on tribal and other cultural groups has largely focused on the way cultural evolution supports more between-group variation than genes. The study of genetic group selection has focused on the evolution of traits that benefit groups at a cost to the individuals that bear them. Cultural evolution does not enhance this form of group selection. However, it can greatly increase the importance of between group competition when within group processes stabilize many different equilibria. Multiple stable equilibria can result from moral norms or institutions backed by systems of reputation, reward, and punishment,[27] from a conformist cultural learning psychology,[28] or from evolved psychological mechanisms that give rise to multiple attractors.[29] When there are multiple equilibria, adaptive processes like selection and payoff biased cultural learning can create and maintain variation among groups. Because cultural adaptive processes are typically much faster than natural selection acting on genetic variation, this process is likely to be more important in cultural evolution than in genetic evolution. It's also important to note that this process is not sensitive to group size. Multiple equilibria and rapid cultural adaptation can sustain variation among very large groups,[30] and thus can account for the high levels of cultural variation among groups.

If groups compete and groups maintained at different equilibria differ in their success in such competition, then selection will lead to the spread of cultural traits that lead to success. Three different forms of competition have been studied. If group equilibria affect extinction rate of groups and if empty habitats are colonized by immigrants mainly drawn from a single surviving group, then traits leading to lower extinction rates will spread.[31] Group beneficial traits can also spread if individuals preferentially imitate their more successful neighbors.[32] For obvious reasons, this process cannot work in a well-mixed population of groups and requires some spatial structure. Finally, the traits present at group beneficial equilibria can increase in frequency if migrants preferentially move from less successful groups to more successful groups and immigrants are assimilated in their new groups.[33] All of these models require that some process causes groups to occasionally shift equilibria, and there are a number of plausible processes that can create such “group mutations”, even in large groups.[34]

Empirical evidence

The empirical evidence that cultural group selection is an important force in human evolution is quite strong.[35] To make a case for cultural group selection we must show (1) that there is appreciable cultural variation between groups, (2) that groups compete, and (3) that cultural variation affects success and failure in intergroup competition. Cultural between-group variation in opinions plausibly related to culture at the level of nations and ethnic groups is more than 10-fold greater than genetic variation.[36]

Variation in institutions is especially important. Institutions are formalized culturally inherited rules that organize our social behavior. These rules institutions prescribe rewards for conformity to the institution and punishments for deviance. Institutions are inherently the properties of groups. A fairly significant number of people have to participate in an institution to make it function. Successful institutions create equilibria within groups, and because many different equilibria are stable, selection between equilibria is possible.

Competition between culturally variable entities is a staple of history. The Romance cultures of Europe and Latin America inherit the institutions of the Roman Empire, such as its legal system, as well as other cultural attributes. The geography of languages preserves histories of competitive expansions of the past, such as the Bantu expansion from western Africa into eastern and southern Africa.[37] Groups of all sorts compete—tribes, business organizations, churches, political parties, modern nations, local governments, and civil society organizations. Group level cultural variation is consequential in these competitions. For example, in many contemporary societies, people invited to play iterated laboratory games of cooperation use altruistic punishment to organize cooperation.[38] However, in other societies, antisocial punishment is common and cooperation remains low in the laboratory. Such variation maps onto real-world economic performance.[39]

Genetic and cultural evolution are tightly coupled. Cultural evolution is driven by human choices, and these choices are often dependent on our genetic make-up. Human cuisines must be adapted to our digestive physiology and its requirement for relatively nutrient dense foods. Many cultural techniques, such as cooking, milling, leaching, and fermentation are devoted to making otherwise inedible products suitable for human consumption. Following E.O. Wilson and Hamilton,[40] many evolutionary students of human behavior have stressed this form of coevolution. On the other hand, culture creates the environments humans live in. Take cuisines again. Unusual diets generate selection on dietary physiology. In some cultures, dairying has led to high frequencies of adult lactase persistence. Lactase is typically down-regulated about the time of weaning in mammals, as there is no natural source of lactose besides milk.[41] Similarly, High Arctic people who eat large quantities of fat from marine mammals and fish are adapted to deal with the high concentrations of polyunsaturated fatty acids in such a diet.[42] More controversially, we have proposed that living in tribes exerted strong selection on innate aspects of our social psychology, what we have called the tribal social instincts hypothesis.[43] For example, humans readily form affectionate social bonds with mates and friends. Using such bonds leads to great benefits from information sharing and many other forms of cooperation. Our normative/institutionalized form of cooperation acts as a form of social selection.[44] For example, laws and informal sanctions handicap the fitness of psychopathic individuals. In effect, humans domesticated themselves in order to reap the benefits of cultural adaptations. Migrants entering a new group are generally subject to the norms and institutions of their host group and are this incentivized to obey them. Often, after two or three generations migrants largely assimilate to their host culture.

The Price equation and cultural group selection

A number of authors have used the Price equation to analyze cultural group selection. For example, Henrich[45] used the Price's equation to illustrate how cultural evolution could generate cooperation in large groups of non-relatives via cultural group (or multi-level) selection whereas the conditions for doing the same given genetic inheritance are much more limited. Bell et al.[46] used the Price's equation to show how much more altruism culture might support in various kinds of human groups than genes given measured amounts of between-group cultural and genetic variation. If these analyses were correct, then it would follow that cultural group selection could be thought of analogous to kin selection, except that "relatedness" would measure the cultural correlations among interacting individuals.

However, the Price equation cannot do full justice to the cultural evolution of human cooperation because its application to structured populations typically assumes weak selection. Strong adaptive forces must be involved if groups that exchange migrants or ideas are to maintain difference in their norms and institutions, as the empirical facts suggest they do. And this means that Hamilton's rule, the Price equation, and the usual machinery used to understand the genetic evolution of behavior cannot be imported to understand cultural group selection.

A model with strong selection

To see why, let's think about a very simple model of the process. A very large population is subdivided into groups of size *n*. Individuals may cooperate within the group to produce a collective good. For example, they may cooperate in the construction of a fish trap or irrigation canal that increases the amount of food available to the group. There are two culturally transmitted norms, labeled red and blue, that affect whether individuals participate. Individuals with the blue norm ("blues") believe that they are obligated to help produce the collective good, and think that others should contribute as well. As a result, they reward others who contribute and don't reward those who don't contribute. Red individuals ("reds) don't contribute and don't reward contributors. This means that blues reward other blues, but do not reward reds. An individual's payoff depends on her behavior and the composition of her group. The more blues that there are in a group the bigger or better the fish trap, and both blues and reds benefit equally. Producing this benefit is costly but it is also normative among blues and they reward individuals who provide the group benefit. When blues are rare, the costs exceed these normative benefits, but when blues are common, the reverse is true. Reds do not contribute and so enjoy the exactly the same benefits produced by the fish trap as the blues.

We formalize these ideas by specifying how the individual payoff of both types depends on group composition. Let x_i be the fraction of individuals in group *i* who are blues The payoff for a red individual in group *i*, labeled *u*, is:

$$u = 1 + gx_i \tag{1}$$

Both reds and blues receive benefits produced by blues in proportion to the frequency of blues. The magnitude of this effect is proportional to the parameter g which measures the increase in benefits of the collective good. The payoff for a blue individual, labeled *v*, is:

$$v = 1 + s(x_i - z) + gx_i \tag{2}$$

where *s* is positive. When blue is rare, the term $s(x_i - z)$ is negative---the individual cost contributing to the collective good exceeds the normative benefits provided by other blues. When blue is common, the term $s(x_i - z)$ is positive---producing the benefits is individually beneficial. The parameter *z* gives the threshold frequency of blues in a group necessary for the normative benefits to exceed the cost. This means that blues have a higher payoff than reds in their group if $x_i > z$. Red individuals do better than blues when $x_i < z$ in their group. The larger the value of the parameter *s*, the larger are the individual costs and benefits compared to the benefits due to the collective good, in this case, the fish trap that accrue to reds and blues equally.

Now, we assume that after social interaction individuals meet a cultural model and observe his or her payoff. With probability *m* the model is drawn from another, randomly chosen, group, and with probability 1− m from the individual's own group. The observer adopts the model's variant with probability

$$\tfrac{1}{2}\big(1 + \beta(w_m - w_o)\big) \tag{3}$$

Where w_o and w_m are the payoffs of the oberver and the cultural model respectively. This means that individuals tend to switch to their cultural model's behavior if the model has a higher payoff. The parameter β controls the strength of this process, called biased transmission. The larger the value of β, the more likely it is that the observer will switch to higher payoff behaviors. Biased transmission increases the frequency of higher payoff behaviors in the populations. Larger values of β creating more rapid adaptive change.

Both strong and weak biases can generate group beneficial behavior but in very different ways. Lehmann et al.[47] derive a condition for blue, the group beneficial variant, to increase in the whole population when rare. In the notation used here this condition is:

$$zs < \frac{r(s + mg)}{\big(1 - r(1 - m)\big)} \tag{4}$$

Where $r = 1/\big(nm(2 - m)\big)$ is the equilibrium relatedness when bias is weak and groups are large. When individuals with the group beneficial variant are rare, they suffer a payoff disadvantage relative to the common variant. The magnitude of this disadvantage is proportional to zs, so the left-hand side is the cost associated with the group beneficial variant when it is rare. The right-hand side is proportional to the relatedness and gives the inclusive fitness benefit associated with the group beneficial variant. This is a version of Hamilton's rule tailored to this particular model that predicts when the frequency of rare blue-types will increase. In Figure 1 we simulate this model under conditions that favor blue.

Figure 1 about here

But a little thought should lead the reader to puzzlement. The direction of evolution in different groups depends on the frequency of red and blue in that group. If relatedness is not zero, there will be some groups in which the frequency of blue is greater than the threshold, z, and so blue will increase in those groups because blue individuals achieve a higher payoff. If selection is strong compared to migration this will lead to different frequencies of blue and red in different groups at equilibrium. How is it possible that we can make a model in which we need only keep track of the change in frequency in the population as a whole?

The answer is *weak selection*. Lehmann et al assume that the selective process generated by biased transmission is so weak that it does not affect the distribution of variants across groups. As selection changes the frequency of blue in the group, the distribution of variants across groups reaches an equilibrium only influenced by migration and drift. Because the payoffs are linear functions of group frequency all you need to calculate the average payoff of an individual is the average number of blues in the group, nr.

Weak selection is a key assumption in the derivation of Hamilton's rule more generally.[48] When the model is applied to the genetic evolution of altruism it is probably a good assumption in many settings. But weak selection cannot capture the dynamics generated by rapid cultural evolution.

Now suppose that payoff bias is strong enough relative to mixing that once either trait is common within a group, it will remain common, and that blue is initially common in some groups. In this case nothing happens even though (4) is satisfied. The blue trait still raises payoffs, and individuals in the groups in which it is common are still disproportionately imitated by individuals in other groups, and relatedness is high. However, unlike weak selection models, producing more emigrants is not enough. The group beneficial trait does not spread to groups in which it is not common, because payoff bias acts strongly against the trait in such groups. Thus, the group beneficial trait remains common in populations in which it is initially common, but cannot spread.

The group beneficial trait can spread, even in very large groups, if the model is modified in one of two ways. First suppose that groups with a higher frequency of the group beneficial trait are less likely to suffer extinctions, and that empty habitats are recolonized by individuals drawn from a single randomly selected group.[49] This assumption is consistent with ethnographic, historical, and archaeological research.[50] Figure 2 shows that the group beneficial trait increases, and relatedness remains high even though the groups are an order of magnitude larger than in the weak bias case. When n = 1000 and payoff-bias is weak, r is very low, and the group beneficial trait is unlikely to spread. Notice that the dynamics of the distribution of frequencies across groups is very different than in the weak bias case—throughout the process strong bias maintains quite different frequencies of the group beneficial trait among groups, and adaptation occurs because groups with a low frequency of the group beneficial variant are more likely to go extinct than groups with a high frequency of the variant. Very similar dynamics result if extinctions are random, but groups with high frequencies of the group beneficial variant are more likely to grow and colonize empty patches, or if groups engage in conflicts with other groups and groups with a higher frequency of the group beneficial variant are likely to be victorious.

The second way to modify the strong bias model is to use a stepping stone population structure so that individuals only imitate models in a small number of neighboring groups.[51] Lehmann et al assume island model migration, and this is sensible given their assumption of weak bias—there is only a modest difference between island and stepping stone models in the weak bias case. However, when bias is strong, the difference is crucial, and cannot be ignored. If the group beneficial trait becomes common in one group, the high payoff causes individuals in neighboring groups to adopt the group beneficial variant, which can tip the neighbors into the basin of attraction of the group beneficial trait. This results in a cascade which spreads the group beneficial trait throughout the population. This process is formally similar to genetic models of the third phase of Wright's shifting balance process,[52] the dynamics of hybrid zones,[53] and early models of reciprocal altruism.[54] The difference is that cultural adaptation can maintain sharp heritable behavioral differences among neighboring human social groups—neighboring ethnolinguistic groups numbering a few thousand people living a few kilometers apart can have mutually unintelligible languages and strikingly different moral systems. Clines on this scale do not seem to occur with genetically transmitted influences on social behavior within other mobile mammal species, because, we believe, selective processes are usually not strong enough relative to migration among local groups.

Thus, the beneficial variant can spread when it is rare in the population as a whole as long as some random nonadaptive process causes it to become common in a single local group. There are at least three plausible processes that can do this. First, errors and sampling variation leads random fluctuations in the frequency of different variants,[55] and this can lead to "peak shifts" for the same reasons as genetic drift. This mechanism depends on sampling variation and should be less effective in large groups. However, the cultural analog of mutation rates are much higher than genetic mutation rates and as a result waiting times for peak shifts can be short even in very large groups.[56]

Second, in random environments, the cultural analog of linkage leading to "genetic draft"[57] can also lead to peak shifts, and the rate at which this occurs does not depend strongly on group size. Linkage in cultural transmission occur when learners acquire two traits from the same person, either because that person is a particularly salient model or because acquiring one trait increases receptivity to a second trait. This leads to correlations between traits analogous to linkage disequilibrium. Then payoff biases that increase the frequency of one trait also tend to increase the frequency of correlated traits, and in a fluctuating environment this leads to random, nonadaptive temporal variation in frequencies that can cause the shift from one basin of attraction to another. Finally, both individual learning and biased transmission depend on environmental cues. One cue will cause an individual to preferentially adopt one variant, while a different cue will cause her to adopt the alternative variant. The cues observed by members of a group may often be highly correlated. For example, the disastrous loss of WWII seems to have shifted Japanese from militaristic moral system to a more pacifistic one.[58] However, such cues often have a strong random component, and as a result will lead to random fluctuations in the frequency of different behaviors. The decisive American victory in the naval Battle of Midway could have easily turned out differently, and if so, the Japanese might have learned a different lesson.

Conclusion

Does Hamilton still rule? Hamilton pioneered the theory of the evolution of cooperation and his 1975 chapter is a pioneering application to human cooperation. All subsequent work in a field takes place in the shadow of the pioneers' contributions or we would not call them pioneers. In this sense, Hamilton still rules. Of course, subsequent workers have made their contributions, some of them rather fundamental, such as Trivers work on reciprocal altruism.[59] Here we have argued that in the case of humans, and perhaps in other cultural species, culture introduces novel evolutionary processes that these must be added to the usual recipe in the Modern Synthesis. In particular we argue that culture introduces selection-like forces that can be much stronger than natural selection on genes. As a result, cultural evolution can generate substantial between group cultural variation in the face of realistic migration rates much more easily than natural selection can maintain genetic difference between groups. Hamilton[60] correctly identified that tribal facies problem that had to be solved. He was not very happy with the gene based inclusive fitness solutions he discussed. He was aware that a cultural explanation was possible but was skeptical that it would prove practical. This was several years before the first formal models of the cultural evolution of cooperation were attempted. Today, the cultural group selection route still remains controversial in some quarters. However, the main alternative is models that assume that individuals achieve mutually beneficial norms by self-interested bargaining, forms of reciprocal altruism.[61] So Hamilton's inclusive fitness rule has not ruled as the solution to tribal facies problem.

The field of cultural evolution only really began to flourish around the time of Hamilton's demise. We wonder what he would have thought of it all if not for his tragic death.

End Notes

[1] Hamilton, 1975

[2] Boyd, 2018; Boyd & Richerson, 2021

[3] Powers & Lehmann, 2017

[4] Bell et al., 2016

[5] Bowles, 2006

[6] Taylor, 1992

[7] Lehmann et al., 2006; Schonmann & Boyd, 2016

[8] Grafen, 1985

[9] Wall et al., 2008

[10] Bell et al., 2016; Jones, 2018

[11] Hamilton, 1975 p. 51

[12] Binmore, n.d.; Boyd, 2018b; Panchanathan & Boyd, 2004; Powers et al., 2016; Powers & Lehmann, 2017; Richerson & Boyd, 2005a

[13] Mathew, 2017

[14] Mathew, 2017; Mathew & Boyd, 2011, 2014

[15] Mathew, 2017

[16] Hamilton, 1975 p. 134

[17] Richerson et al., 2016

[18] E.g. (Pinker, 2012)

[19] Gintis et al., 2001; Jordan et al., 2016

[20] Hooper et al., 2010

[21] Bhui et al., 2019; Raihani & Bshary, 2019

[22] Campbell, 1965

[23] Cavalli-Sforza & Feldman, 1973

[24] Boyd & Richerson, 1985

[25] Perreault, 2012

[26] Vogel et al., 2009

[27] Boyd et al., 2011a; Boyd & Richerson, 2009a

[28] Boyd & Richerson, 1982

[29] Henrich & Boyd, 2016

[30] Boyd et al., 2011b

[31] Boyd et al., 2011b; Boyd & Richerson, 1990

[32] Boyd & Richerson, 1992

[33] Boyd & Richerson, 2009b

[34] Arieli & Young, 2016; R Boyd et al., 2011b

[35] Richerson et al., 2016

[36] Bell et al., 2016; Hadley & Mathew, 2020

[37] Currie & Mace, 2009

[38] Fehr & Gachter, 2001

[39] Herrmann et al., 2008

[40] Hamilton, 1975; Wilson, 1978

[41] Ségurel & Bon, 2017

[42] Fumagalli et al., 2015

[43] Richerson & Boyd, 2005b

[44] Boehm, 2012

[45] Henrich, 2004

[46] Bell et al., 2016

[47] Lehmann et al., 2008

[48] Grafen, 1985

[49] Boyd & Richerson, 1990

[50] Bowles, 2009; Keeley, 1997; Soltis et al., 1995

[51] Boyd & Richerson, 2002

References

Arieli, I., & Young, H. P. (2016). Stochastic Learning Dynamics and Speed of Convergence in Population Games. *Econometrica*, *84*(2), 627–676.

Barton, N. H. (1979). The dynamics of hybrid zones. *Heredity*, *43*(3), 341–359.

Bell, A. V, Richerson, P. J., & McElreath, R. (2016). Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceeding of the National Academy of Sciences USA*, *106*(42), 17671–17674.

Bhui, R., Chudek, M., & Henrich, J. (2019). How exploitation launched human cooperation. *Behavioral Ecology and Sociobiology*, *73*(6). https://doi.org/10.1007/s00265-019-2667-y

Binmore, K. (2005). *Natural Justice*. Oxford, Oxford University Press.

Boehm, C. (2012). Moral origins: the evolution of virtue, altruism, and shame. In *Choice Reviews Online* (Vol. 50, Issue 02). Basic Books,. https://doi.org/10.5860/choice.50-0864

Boorman, S. A., & Levitt, P. R. (1973). A frequency-dependent natural selection model for the evolution of social cooperation networks. *Proceedings of the National Academy of Sciences, 70*(1), 187–189.

Boorman, S. A., & Levitt, P. R. (1980). *The Genetics of Altruism*. Academic Press.

Bowles, S. (2006). Group competition, reproductive leveling and the evolution of human altruism. *Science, 314*, 1569–1572.

Bowles, S. (2009). Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science, 324*(5932), 1293–1298.

Boyd, R. (2018a). *A Different Kind of Animal: How Culture Transformed Our Species*. Princeton

Boyd, R, & Richerson, P. J. (1982). Cultural transmission and the evolution of cooperative behavior. *Human Ecology, 10,* 325–351.

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

Boyd, R, & Richerson, P. J. (1990). Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology*, *145*(3), 331–342.

Boyd, R, & Richerson, P. J. (2002). Group beneficial norms can spread rapidly in a structured population. *Journal of Theoretical Biology*, *215*, 287–296.

Boyd, R, & Richerson, P. J. (2009a). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3281–3288.

Boyd, R, & Richerson, P. J. (2009b). Voting with your feet: Payoff biased migration and the evolution of group beneficial behavior. *Journal of Theoretical Biology*, *257*, 331–339.

Boyd, R, & Richerson, P. J. (2021). Large-scale cooperation in small-scale foraging societies. *EcoEvoRxiv*, *10*(32942/osf.io/fxwbr).

Boyd, R, Richerson, P. J., & Henrich, J. (2011). Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. *Behavioral Ecology and Sociobiology*, *65*, 431–444.

Boyd, Robert, & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, *13*(3), 171–195. https://doi.org/10.1016/0162-3095(92)90032-Y

Campbell, D. T. (1965). *Variation and selective retention in socio-cultural evolution. Social Change in Developing Areas: A Reinterpretation of Evolutionary Theory* (H. R. Barringer, G. I. Blanksten, & R. W. Mack (eds.)). Schenkman Publishing Company.

Cavalli-Sforza, L. L., & Feldman, M. W. (1973). Cultural versus biological inheritance: Phenotypic transmission from parents to children (A theory of the effect of parental phenotypes on children's phenotypes. *American Journal of Human Genetics*, *25*, 618–637.

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

Currie, T. E., & Mace, R. (2009). Political complexity predicts the spread of ethnolinguistic groups. *Proceedings of the National Academy of Sciences*, *106*(18), 7339–7344.

Dower, J. W. (1999). *Embracing Defeat: Japan in the Wake of World War II*. WW Norton & Company.

Fehr, E., & Gachter, S. (2001). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140.

Gavrilets, S. (1996). On phase three of the shifting-balance theory. *Evolution*, *50*(3), 1034–1041.

Gillespie, J. H. (2000a). Genetic drift in an infinite population: The pseudohitchhiking model. *Genetics*, *155*(2), 909–919.

Gillespie, J. H. (2000b). Is the population size of a species relevant to its evolution? *Evolution*, *55*(11), 2161–2169.

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly Signaling and Cooperation. *Journal of Theoretical Biology*, *213*, 103–119.

Grafen, A. (1985). A geometric view of relatedness. *Oxford Surveys in Evolutionary Biology*, *2*(2), 28–89.

Hadley, C., & Mathew, S. (2020). *Human large-scale cooperation as a product of competition between cultural groups*. *11*, 72–80.

Hamilton, W. D. (1975). *The innate social aptitudes of man, Biosocial Anthropology* (R. Fox (ed.)). Malby Press.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, *53*(1), 3–35.

Henrich, J., & Boyd, R. (2016). How evolved psychological mechanisms empower cultural group selection. *Behavioral and Brain Sciences*, *19,* 68.

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science, 319*, 1362–1367.

Hooper, P. L., Kaplan, H. S., & Boone, J. L. (2010). A theory of leadership in human cooperative groups. *Journal of Theoretical Biology*, *265*, 633–646.

Jones, D. (2018). Kin selection and ethnic group selection. *Evolution and Human Behavior*, *39*, 9–18.

Jordan, J. J., Hoffman, M., Bloom, P., & Rand, D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, *530,* 473–476.

Keeley, L. (1997). *War Before Civilization: The Myth of the Peaceful Savage*. Oxford Univ Press.

Kreindler, G. E., & Young, H. P. (2013). Fast convergence in evolutionary equilibrium selection. *Games and Economic Behavior*, *80,* 39–67.

Lehmann, L., Feldman, M. W., & Foster, K. (2008). Cultural transmission can inhibit the evolution of altruistic helping. *The American Naturalist*, *172,* 12–24.

Lehmann, L., Perrin, N., & Rouset, F. (2006). Population demography and the evolution of helping behaviors. *Evolution*, *60*, 1137–1151.

Mathew, S. (2017). How the second-order free rider problem is solved in a small-scale society. *American Economic Review*, *107*(5), 578–581. https://doi.org/10.1257/aer.p20171090

Mathew, S., & Boyd, R. (2011). Punishment sustains large-scale cooperation in prestate warfare. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(28), 11375–11380. https://doi.org/10.1073/pnas.1105604108

Mathew, S., & Boyd, R. (2014). The cost of cowardice: Punitive sentiments towards free riders in turkana raids. *Evolution and Human Behavior*, *35*(1), 58–64. https://doi.org/10.1016/j.evolhumbehav.2013.10.001

Matteo Fumagalli, M., Moltke, I., Grarup, N., Racimo, F., Bjerregaard, Marit, P., Jørgensen, E., Korneliussen, T. S., Gerbault, P., Skotte, L., Linneberg, A., Christensen, 1C, Brandslund, I., Jørgensen, T., Huerta-Sánchez, E., Schmidt, E. B., Pedersen, O., & Hansen, TAlbrechtsen, A Nielsen, R. (2015). Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science*, *349*(6254), 1343–1347.

Neiman, F. D. (1995). Stylistic variation in evolutionary perspective: Inferences from decorative diversity and interassemblage distance in Illinois Woodland ceramic assemblages. *American Antiquity*, *60*, 7–36.

Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, *432*, 499–502.

Perreault, C. (2012). The pace of cultural evolution. *PLOS ONE*, *7*(9), 45150.

Pinker, S. (2012). The false allure of group selection. *Edge*.

Powers, S. T., & Lehmann, L. (2017). When is bigger better? The effects of group size on the evolution of helping behaviours. *Biological Reviews*, *92*, 902–920.

Powers, S. T., Schaik, C. P., & Lehmann, L. (2016). How institutions shaped the last major evolutionary transition to large-scale human societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1687), 20150098.

Raihani, N., & Bshary, R. (2019). Punishment: one tool, many uses. *Evolutionary Human Sciences*. https://doi.org/10.1017/ehs.2019.12

Richerson, P., Baldini, R., Bell, A., Demps, K., Frost, K., Hillis, V., Mathew, S., Newton, E., Narr, N., Newson, L., Ross, C., Smaldino, P., Waring, T., & Zefferman, M. (2016). Cultural group selection plays an essential role in explaining human cooperation: A sketch of the evidence, together with commentaries and authors' response. *Behavioral and Brain Sciences*, *39*(e 30), 1–68.

Richerson, P. J., & Boyd, R. (2005). *Not By Genes Alone: How Culture Transformed Human Evolution*. Chicago, University of Chicago Press.

Schonmann, R. H., & Boyd, R. (2016). A simple rule for the evolution of contingent cooperation in large groups. *Philosophical Transactions of the Royal Society (B*, *317*. https://doi.org/10.1098/rstb.2015.0099

Ségurel, L., & Bon, C. (2017). On the Evolution of Lactase Persistence in Humans. *Annual Review of Genomics and Human Genetics*, *18*, 297–319. https://doi.org/10.1146/annurev-genom-091416-035340

Shennan, S. (2001). Demography and cultural innovation: A model and its implications for the emergence of human culture. *Cambridge Archaeological Journal*, *11*(1), 5–16.

Soltis, J., Boyd, R., & Richerson, P. J. (1995). Can group-functional behaviors evolve by cultural group election? An empirical test. *Current Anthropology*, *36*(3), 473–494.

Taylor, P. D. (1992). Inclusive fitness in a homogeneous environment. *Proceedings of the Royal Society of London B*, *249*, 299–302.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*(1), 35–57.

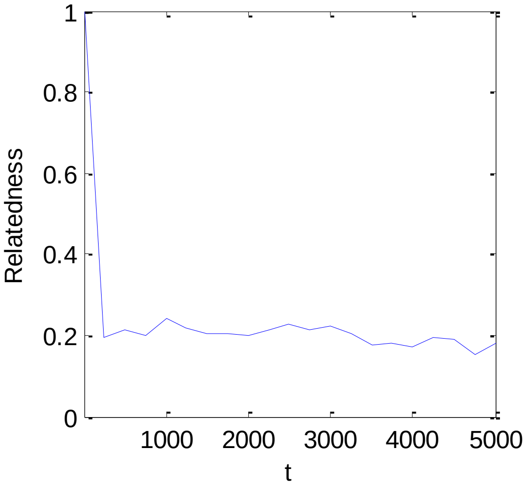
Vogel, V., Pedersen, J. S., D’Ettorre, P., Lehmann, L., & Keller, L. (2009). Dynamics and genetic structure of argentine ant supercolonies in their native range. *Evolution, 63*(6), 1627–1639.

Wall, J. D., Cox, M. P., Mendez, F. L., Woerner, A., Severson, T., & Hammer, M. F. (2008). A novel DNA sequence database for analyzing human demographic history. *Genome Research*, *18,* 1354–1361.

Young, H. P. (1995). The evolution of social norms. *Annual Review of Economics*, *7,* 359–387.

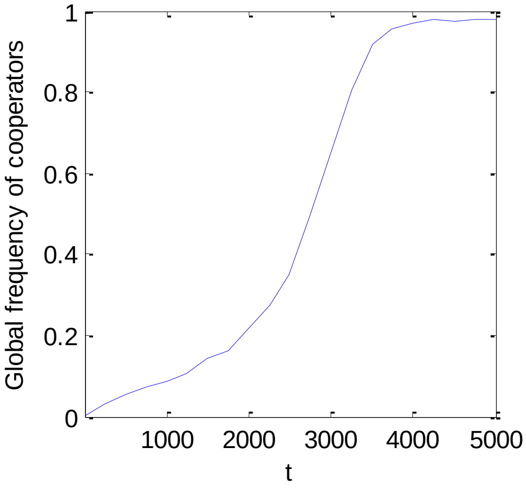
Figures

(a)



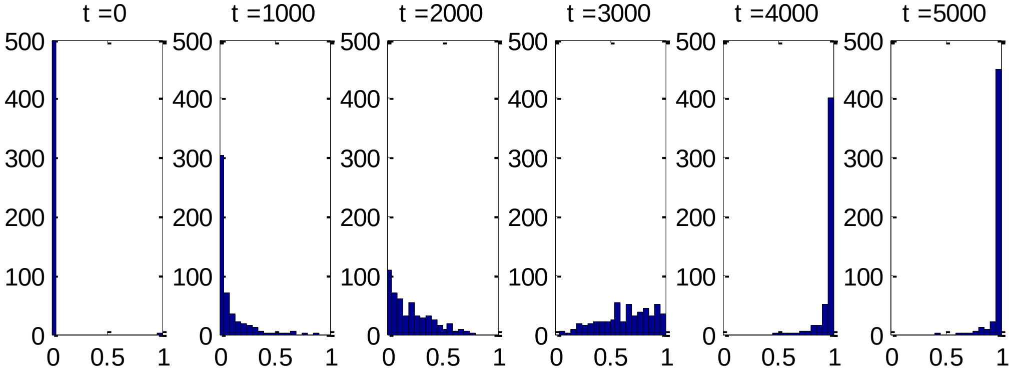
1
0.8
0.6
0.4
0.2
0
Relatedness
1000
2000
3000
4000
5000
t

(b)



1
0.8
0.6
0.4
0.2
0
Global frequency of cooperators
1000
2000
3000
4000
5000
t

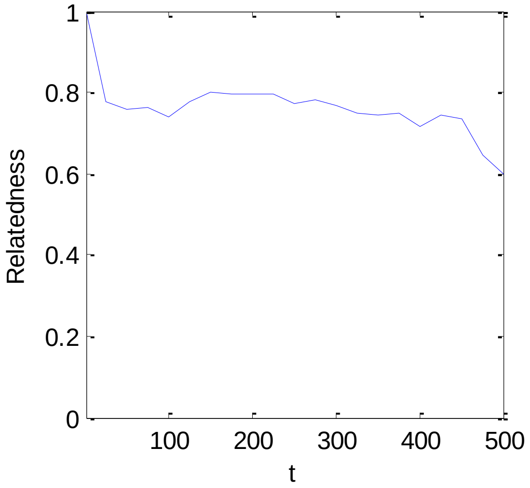
(c)



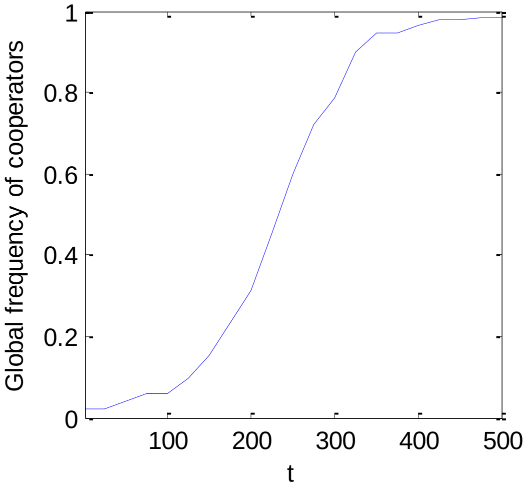
t =0
t =1000
t =2000
t =3000
t =4000
t =5000
500
400
300
200
100
0
0
0.5
1

Figure 1. The dynamics of the group beneficial trait when payoff bias is weak (n = 100, m = 0.02, s = 0.2, g = 1, $z = 0.25$, b = 0.05). There are 500 groups. Initially, there is one group in which the frequency of the group beneficial variant is one in one group, and zero in all other groups. (a) shows that relatedness within groups quickly attains the predicted equilibrium value (≈0.2 for these parameter values), and (b) shows that because the inclusive fitness benefits exceed the costs, the group beneficial trait increases in frequency. (c) Gives the distribution of frequencies across groups as the group beneficial trait increases. Individuals with the group beneficial variant rapidly diffuse throughout the population, and then the distribution of frequencies results from the interplay of migration and common descent. Adaptation can be understood as responding to the average over the entire distribution.

(a)


Relatedness
1
0.8
0.6
0.4
0.2
0
100
200
300
400
500
t

(b)


Global frequency of cooperators
1
0.8
0.6
0.4
0.2
0
100
200
300
400
500
t

(c)

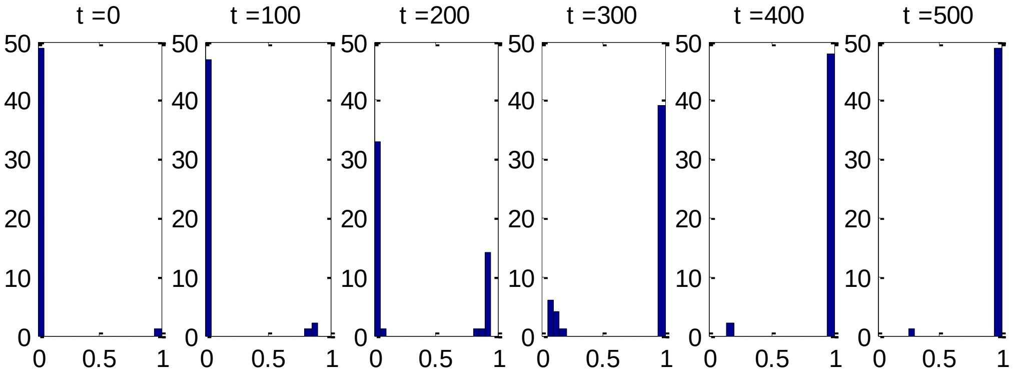

t =0
t =100
t =200
t =300
t =400
t =500
50
40
30
20
10
0
0
0.5
1

Figure 2. The dynamics of the group beneficial trait when bias is strong and the group beneficial trait lowers extinction rates (n = 1000, m = 0.02, s = 0.2, g = 0, $z = 0.5$, □ = 1.0). There are 50 groups. The probability of extinction in group i each time period is $\varepsilon(1 - x_i)$ where x_i is the frequency of the group beneficial trait in group i and □ = 0.015, a value that when combined with the distribution of frequencies yields extinction rates roughly consistent with those observed in tribal societies (Soltis et al. 1995) assuming simulation time periods of one year. Empty habitats are recolonized by immigrants from a single surviving group. Initially, there is one group in which the frequency of the group beneficial variant is one, and zero in all other groups. (a) Relatedness within groups quickly reaches an equilibrium value of about 0.8 even though groups are quite large because strong bias maintains the group beneficial norm at either a high or low frequency in every group. Here relatedness is the extent to which an individual's behavior predicts the behavior of others in her group, and is mainly not the result of common descent. (b) The group beneficial trait spreads because such groups are much less likely to become extinct. (c) Gives the distribution of frequencies across groups as the group beneficial trait increases. Throughout the process, strong bias maintains groups at strongly different frequencies, and adaptation occurs because groups with a low frequency of the group beneficial variant are more likely to go extinct than groups with a high frequency of the variant.