

## Original Article

## Hunter–Gatherer population structure and the evolution of contingent cooperation

Robert Boyd <sup>a,b,\*</sup>, Roberto H. Schonmann <sup>c</sup>, Renato Vicente <sup>d</sup><sup>a</sup> School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA<sup>b</sup> Santa Fe Institute, Santa Fe, NM, USA<sup>c</sup> Department of Mathematics, University of California at Los Angeles, CA 90095, USA<sup>d</sup> Department of Applied Mathematics, Instituto de Matemática e Estatística, Universidade de São Paulo, 05508–090, São Paulo-SP, Brazil

## ARTICLE INFO

## Article history:

Initial receipt 5 June 2013

Final revision received 2 February 2014

## Keywords:

Cooperation

Hunter–Gatherers

Population structure

## ABSTRACT

Unlike other vertebrates, humans cooperate in large groups with unrelated individuals. Many authors have argued that the evolution of such cooperation has resulted from reciprocity and other forms of contingent cooperation. This argument is not well supported by existing theory. The theory of contingent cooperation in pairs is well developed: reciprocating strategies are stable when common, and can increase when rare as long as population structure leads to modest levels of relatedness. In larger groups, however, it is not clear whether contingent cooperation can increase when rare. Existing work suggests that contingent strategies cannot increase unless relatedness is high, but depends on unrealistic assumptions about the effects of population structure. Here we develop and analyze a model incorporating a two level population structure that captures important features of human hunter–gatherer societies. This model suggests that previous work underestimates the range of conditions under which contingent cooperation can evolve, but also predicts that cooperation will not evolve unless (1) social groups are small, and (2) the relatedness within ethnolinguistic groups is at the high end of the range of empirical estimates.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

Humans cooperate in large groups of unrelated individuals. This is obviously true in modern societies in which complex institutions allow the mobilization of millions of individuals in cooperative enterprises. However, ethnographic and historical evidence indicates that foragers and people living in other acephalous societies that lack such institutions often cooperate in groups of hundreds of individuals. Examples include warfare (Mathew & Boyd, 2011), the construction of capital facilities (Swezey & Heizer, 1977), and large scale cooperative foraging (Nuligak, 1971). The received view is that such cooperation is explained by reciprocity and other forms of contingent behavior.

This view is not theoretically well-supported. It is true that contingent strategies can stabilize cooperation and the punishment of noncooperators. Because people can recognize individuals and remember past behavior, selection can favor a psychology that motivates actors to help only those who have helped them in the past, or punish those who don't cooperate in mutually beneficial activities. If, in the long run, benefits of sustained cooperation exceed the short term benefits of defection, then such a psychology can be evolutionarily stable. However, this is not a sufficient account of human cooperation because it does not explain how strategies supporting contingent cooperation can increase when rare. This

lacuna is important for two reasons: first contingent cooperation appears to be very rare among primates (Clutton-Brock, 2009), and thus it is very likely that the ancestral condition in the human lineage is noncooperative. Second, the combination of repeated interactions and contingent rewards and punishment can stabilize virtually any behavior—prosocial helping, ruthless spite, and everything in between (Boyd & Richerson, 1992). This means it is not enough to explain the stability of contingent cooperation. We must also explain why mutually beneficial strategies are likely evolutionary outcomes. This is a problem because rare contingent cooperators are altruists who pay the cost of initial cooperation but do not gain any long run benefit.

In theory low levels of relatedness can allow a psychology supporting cooperation between pairs of individuals to increase when rare (e.g. Axelrod & Hamilton, 1981; McElreath & Boyd, 2007). If interactions are repeated many times, the benefits to cooperation can be very large, and as a result cooperators can increase even if they have only a small chance of interacting with another cooperator. Population structure often leads to low but positive background levels of relatedness, and this provides a possible explanation for the evolution of pairwise reciprocity. We do not know if this is the correct explanation for the evolution of pairwise reciprocity in humans. First, it would seem to apply to many group living vertebrate species, yet reciprocity is very rare. Second, pairwise reciprocity in humans is frequently regulated by social norms enforced by third parties, a mechanism usually invoked only to explain large scale cooperation (Mathew, Boyd, & vanVeelen, 2013).

\* Corresponding author.

E-mail address: [rboyd3@asu.edu](mailto:rboyd3@asu.edu) (R. Boyd).

It is not clear whether low levels of genetic assortment resulting from population structure can play a similar role in the evolution of contingent cooperation in larger groups. Genetic data provide estimates of relatedness within human foraging groups (Bowles, 2006, Langergraber et al., 2011). However knowing relatedness and the costs and benefits of alternative behaviors is not enough to predict evolutionary outcomes. Cooperative strategies can increase when they have higher expected fitness than noncooperative ones. Because contingent cooperation leads to non-linear fitness functions, knowing  $r$ , the expected frequency of rare cooperators in a group with the focal cooperator, is not enough to calculate expected fitness of a focal cooperator or noncooperator (van Veelen, 2009; Schonmann, Vicente, & Caticha, 2013). Instead, you have to know the conditional probability of each different group composition given the strategy of the focal. (Using the version of Hamilton's rule advocated by Gardner, West, and Wild (2011) requires exactly the same probability distributions to compute the multiple regression coefficients which replace marginal fitness effects). This probability distribution can be calculated from empirical estimates of  $r$  only if assumptions are made about how population structure gives rise to genetic assortment. Previous work (Boyd & Richerson, 1988; Boyd, Gintis, & Bowles, 2010) assumed that groups are formed by sampling individuals with a constant relatedness to each other. This assumption is not plausible because the biological processes that generate relatedness lead to interdependencies so that knowing that two individuals share a gene by common descent increases the probability that other members of the group also share that gene by common descent.

Here we derive a simple recursive expression for this probability distribution for a population structure that captures important features of some human foraging populations. Like other primates, human foragers live in residential groups. These “bands” typically number between 10 and 100 people (Hill et al., 2011), but can sometimes be larger in particularly provident environments. Unlike other primates, however, there is much peaceful social interaction between members of different bands, and people frequently change bands (Hill et al., 2011). Many authors believe that forager populations are usually subdivided into ethnolinguistic groups numbering from several hundred to several thousand people in which individuals are identified by language and other culturally transmitted markers of group membership, and that rates of social interaction and rates of marriage between members of different ethnolinguistic groups are much lower than between members of the same ethnolinguistic group (Tindale, 1974). Other authors (Berndt, 1959) have argued that there are no sharp linguistic or cultural boundaries. The question is difficult to resolve because well-studied modern hunter-gatherers usually live among agricultural populations, and thus we must rely on historical and early ethnographic materials, especially from Australia and North America. In this paper, we assume a two level structure with bands nested inside of ethnolinguistic groups. Elsewhere we will present a model that is better suited to a world in which bounded ethnolinguistic groups do not exist.

To create an analytically tractable model, we assume that ethnolinguistic groups are linked through island model migration meaning that each ethnolinguistic group draws immigrants from all other groups with equal probability. Assuming that migration rates are modest and that selection is weak, we derive an expression describing the pattern of genetic assortment among ethnolinguistic groups. We then assume that social interaction takes place in bands whose members are randomly sampled from the ethnolinguistic group. With these assumptions we derive a very simple rule for the probability distribution of numbers of cooperators across bands. We will show that this simplified model provides a good approximation to more realistic calculations which assume large but limited gene flow among bands within ethnolinguistic groups.

We then use this rule to determine when contingent cooperation and punishment can increase when rare, and, given that they can, the frequency of cooperators at equilibrium. This analysis suggests that

tolerant strategies that cooperate when a minority of others in their band also cooperate can increase when rare if relatedness within ethnolinguistic groups conforms to the higher estimates reported in Bowles (2006) but not the lower estimates reported in Langergraber et al. (2011). Such tolerant strategies evolve to a mixed equilibrium at which a minority of cooperators coexists with a majority of non-cooperators. Less tolerant strategies that require most other individuals in the band to cooperate are very unlikely to increase when rare. These results suggest that observed high frequency cooperation is better explained by contingent punishment of non-cooperators than contingent cooperation.

## 2. Population structure and assortment

Consider a population split into a large number of ethnolinguistic groups that are maintained close to size  $n_e$  by density dependent population regulation. We assume that  $n_e$  ranges between several hundred and a few thousand individuals. Each generation a fraction  $m$  of each ethnolinguistic group is replaced by immigrants drawn at random from the population as a whole. We assume that inter-ethnic migration rates are less than 10%. Ethnolinguistic groups are subdivided into bands of size  $n_e$  (that can be between 10 and 50) in which social interaction occurs. Each generation band members are drawn at random from the ethnolinguistic group. There are two genotypes that affect social behavior, a cooperative type labeled A, and a noncooperative type, N. The reproductive success of each individual will depend on her/his own genotype and on the genotypes of the other members of her/his band. The fitness (i.e., the expected number of offspring that reach adult age) of an individual of type A, or N, in a band with a total number  $k$  of types A, is given by

$$w_k^A = 1 + \delta v_k^A, w_k^N = 1 + \delta v_k^N, \quad (1)$$

with the conventions  $v_0^N = 0$  and  $\delta > 0$ . We will refer to  $v_k^A$  and  $v_k^N$  as the marginal fitnesses.

We assume that selection is weak so that  $\delta$  is small. This means that selection changes the frequency of types more slowly than drift and migration change the composition of groups, and, therefore, that the distribution of types across groups can be calculated assuming there is no selection. We also assume that there is enough elasticity in group size that natural selection increases the frequency of the type with higher expected fitness across groups. There are at least two ways that this can happen. First, groups can grow by budding (Goodnight, 1992; Gardner & West, 2006; Schonmann et al., 2013) so that descendants do not compete. Second, density dependent regulation of the size of ethnolinguistic groups maintains their size close to  $n_e$ , but is elastic enough to accommodate group average fitness larger than 1. This requires that the absolute value of the effect of a change in group size on the fitness of individuals in the group is much larger than  $\delta$  but smaller than  $m$  (Schonmann et al., in prep). As the effect of a change in group size on fitness increases in comparison to  $m$ , it becomes more difficult for altruistic strategies to increase, and in the limit the model produces the same effects as the inelastic island model (Taylor, 1992) which assumes fixed group sizes. In this model, groups make equal contributions to the population independent of the number of A types, and thus altruistic behavior cannot evolve. This means that the results presented below represent a best case for the evolution of contingent cooperation.

With these assumptions the frequency of A types in the population will increase when they have higher expected fitness across groups, or

$$\sum_{k=1}^n \Pr(k|A, p) v_k^A > \sum_{k=0}^{n-1} \Pr(k|N, p) v_k^N, \quad (2)$$

where  $\Pr(k|A, p)$  and  $\Pr(k|N, p)$  are the conditional probabilities that focal A and N types find themselves in a band with  $k$  A types. Since  $\delta$  is small, these probabilities can be calculated assuming that the frequency of A is fixed at  $p$  and the distribution of A and N types across ethnolinguistic groups is determined only by drift and migration.

Due to limited migration, members of the same ethnolinguistic group will be related. Because selection is weak, ethnolinguistic groups are large, and migration rates among ethnolinguistic groups are low, the relatedness within ethnolinguistic groups,  $r_e$ , is approximately

$$r_e = \frac{1}{1 + 2n_e m}. \tag{3}$$

(e.g. Hartl and Clark, 2007). Since bands are formed randomly this is also the relatedness among pairs of band members. This means that the expected number of A types in a band containing a focal A type is  $1 + (n - 1)r_e$  because the focal is A and a fraction  $r_e$  of the other  $n - 1$  individuals are also A on average. However, knowing the expected number of A types is not enough to calculate expected fitnesses unless  $w_k^A$  and  $w_k^N$  are linear functions of  $k$ . For example, in the linear public goods game  $v_k^A = kb/n - c$ , where  $b/n$  is the benefit to each individual produced by a cooperative act that costs  $c$ . Then the expected marginal fitness of an A type individual is  $b(1 + (n - 1)r_e)/n - c$ . However contingent strategies typically cause fitness to be a highly nonlinear function of  $k$ . For example, an  $n$ -person analog of tit-for-tat is “cooperate on the first interaction, and continue to cooperate if  $\theta$  or more individuals cooperated on the previous interaction”. If  $k < \theta$  cooperation ceases after the first interaction, but if  $k \geq \theta$  it is sustained for many periods, and this leads to a fitness function with a sharp kink at  $\theta$ . Just knowing the average value of  $k$  is not enough; you also need to know the probability that  $k \geq \theta$ . For such nonlinear fitness functions, the complete conditional probability distributions  $\Pr(k|A, p)$  and  $\Pr(k|N, p)$  are needed to calculate expected fitnesses.

To calculate these distributions we assume that  $n_e$  is large,  $m$  is small and  $n_e m = \ell$ . Then the distribution of the frequency of A types in the ethnolinguistic group of a focal cooperator is approximately given by a beta probability density with parameters  $\ell p + 1$  and  $\ell(1 - p)$ . This follows from the fact that for the island model, the probability density for the fraction  $x$  of cooperators in a randomly chosen group is  $f(x) = \text{beta}(x|\ell p, \ell q) = Cx^{\ell p - 1}(1 - x)^{\ell q - 1}$ , where  $q = 1 - p$  and  $C$  is a normalization constant (Wright, 1931; Crow & Kimura, 1970; Wakeley, 2003, Schonmann et al in prep). The joint density that the focal is an A type and the fraction of A types in the ethnolinguistic group is  $x$  is  $xf(x)$ . Therefore, the conditional probability density that the focal is in an ethnolinguistic group with a fraction  $x$  of A types, given that the focal is an A type is proportional to  $xf(x) = Cx^{\ell p}(1 - x)^{\ell q - 1}$  and hence it is  $\text{beta}(x|\ell p + 1, \ell q)$ .

Since bands are formed by randomly selecting individuals from their ethnic group, the conditional probability that there are  $k - 1$  A types among the other  $n - 1$  individuals in a band of a focal A type given the frequency of A in the focal's ethnolinguistic group is  $x$  is given by a binomial probability distribution with parameters  $x$  and  $n - 1$ . Thus  $\Pr(k|A, p)$  averaged over all ethnolinguistic groups is given by

$$\Pr(k|A, p) = \int_0^1 \text{bin}(k-1|n-1, x) \text{beta}(x|\ell p + 1, \ell q) dx, \tag{4}$$

where  $\text{bin}(k - 1|n - 1, x)$  is the binomial probability of drawing  $k - 1$  hits each with probability  $x$  in  $n - 1$  trials and  $\text{beta}(x|\ell p + 1, \ell q)$  is the probability distribution of  $x$  over ethnolinguistic groups containing a focal A type. Then substituting

the expressions for the binomial distribution and the beta density yields

$$\Pr(k|A, p) = \int_0^1 \binom{n-1}{k-1} x^{k-1} (1-x)^{n-k} \frac{\Gamma(\ell + 1)}{\Gamma(\ell p + 1)\Gamma(\ell, q)} x^{\ell p} (1-x)^{\ell q - 1} dx \tag{5}$$

Now collect terms that involve  $x$

$$\Pr(k|A, p) = \frac{(n-1)!}{(k-1)!(n-k)!} \frac{\Gamma(\ell + 1)}{\Gamma(\ell p + 1)\Gamma(\ell, q)} \int_0^1 x^{\ell p + k - 1} (1-x)^{\ell q - 1 + n - k} dx, \tag{6}$$

where  $\Gamma$  is the gamma function. Then, using the fact that

$$\int_0^1 dx x^{\alpha - 1} (1-x)^{\beta - 1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}, \tag{7}$$

we can write (6) as

$$\Pr(k|A, p) = \frac{(n-1)!}{(k-1)!(n-k)!} \frac{\Gamma(\ell + 1)}{\Gamma(\ell p + 1)\Gamma(\ell, q)} \frac{\Gamma(\ell p + k)\Gamma(\ell q + n - k)}{\Gamma(\ell + n)}. \tag{8}$$

Then using the identity  $\Gamma(z) = (z - 1)\Gamma(z - 1)$  and setting  $k = 1$ , we obtain

$$\begin{aligned} \Pr(1|A, p) &= \frac{\Gamma(\ell + 1)}{\Gamma(\ell + n)} \frac{\Gamma(\ell q + n - 1)}{\Gamma(\ell, q)} \\ &= \frac{(\ell q + n - 2)(\ell q + n - 3) \dots (\ell, q)}{(\ell + n - 1)(\ell + n - 2) \dots (\ell + 1)}. \end{aligned} \tag{9}$$

And, for  $k = 2, \dots, n$ , we obtain

$$\frac{\Pr(k|A, p)}{\Pr(k-1|A, p)} = \frac{(k-2)!(n-k+1)!}{(k-1)!(n-k)!} \frac{\Gamma(\ell p + k)\Gamma(\ell q + n - k)}{\Gamma(\ell p + k - 1)\Gamma(\ell q + n - k + 1)}, \tag{10}$$

which, using again the fact that  $\Gamma(z) = (z - 1)\Gamma(z - 1)$  simplifies to the recursion

$$\Pr(k|A, p) = \Pr(k-1|A, p) \frac{(n-k+1)(\ell p + k - 1)}{(k-1)(\ell q + n - k)}. \tag{11}$$

The probability that the focal is the only A type in its group,  $\Pr(1|A, p)$ , can be calculated using (9). Then the conditional probabilities that there are  $k = 2, \dots, n$  A types  $\Pr(k|A, p)$  can be calculated very easily using (11), for example, a spreadsheet.

To compute the expected fitness of N types, we need to calculate the conditional probability that N types find themselves in a group with  $k$  A types,  $\Pr(k|N, p)$ . However, since the labels A and N are arbitrary it follows that

$$\Pr(k|N, p) = \Pr(n-k|A, q), \tag{12}$$

for  $k = 0, \dots, n - 1$ , and thus we can use (9) and (11) to calculate the expected fitnesses of both A and N types.

### 2.1. When can contingent cooperation invade?

Now consider a population in which A types are rare. If the focal is type N, virtually everyone in her/his ethnolinguistic group is type N, and her/his band will consist of no A types and  $n$  N types. Since the marginal fitness of an N type in a band of all N types is zero, the expected fitness of N types is 1. But if the focal is type A, then the expected fraction of types A in her/his ethnolinguistic group is close to  $r_e = 1/(1 + 2n_e m)$ , and the probability that she/he is in a group with a total of  $k$  types A is

given by  $\Pr(k|A) = \lim_{p \rightarrow 0} \Pr(k|A, p)$ . When  $p \rightarrow 0$ , (9) and (11) simplify to become

$$\Pr(1|A) = \frac{\ell}{n + \ell - 1}, \Pr(k|A) = \Pr(k-1|A) \left( \frac{n-k+1}{n-k+\ell} \right). \quad (13)$$

Types A will proliferate if their expected fitness is greater than that of types N, which is 1, or in terms of marginal fitnesses,

$$\sum_k \Pr(k|A) v_k^A > 0. \quad (14)$$

We consider fitness functions that represent two different types of contingent social behavior. The iterated public goods game captures the payoffs that result from contingent cooperation (Joshi, 1987; Boyd & Richerson, 1988; Doebeli & Hauert, 2005). Interactions last  $T$  periods on average. Each period individuals can either cooperate (C) or defect (D). Each person cooperating increases the fitness of members of its group by  $b/n$  at a cost  $c$ . Defection has no effect on fitness. Behavior is controlled by one of two strategies. A types cooperate during the first period, and during subsequent periods they cooperate if  $\theta$  or more group members cooperated during

the previous period. Otherwise they defect. N types always defect. Then

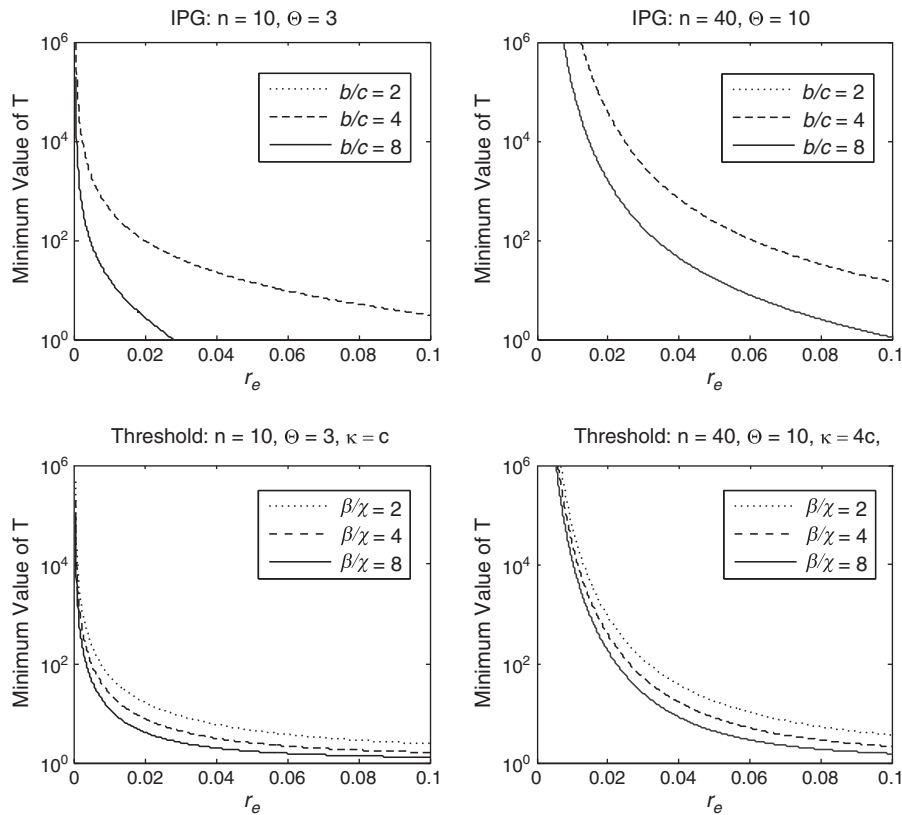
$$v_k^A = \begin{cases} \left( \frac{kb}{n} - c \right) & \text{if } k < \theta, \\ T \left( \frac{kb}{n} - c \right) & \text{if } k \geq \theta, \end{cases} \quad (15)$$

and

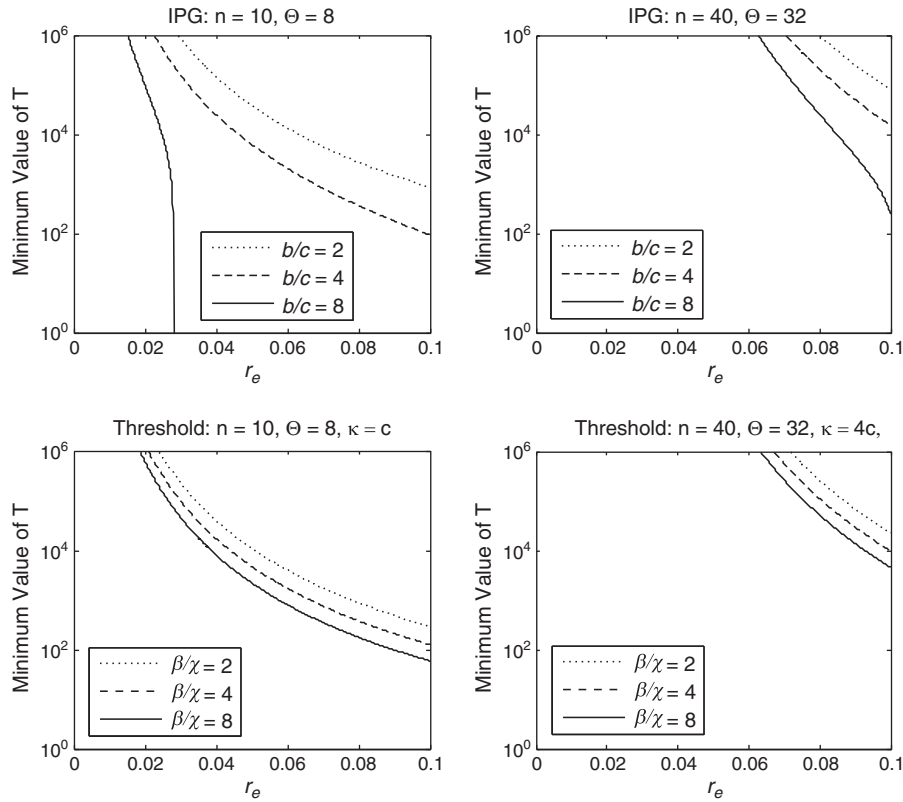
$$v_k^N = \begin{cases} \left( \frac{kb}{n} \right) & \text{if } k < \theta, \\ T \left( \frac{kb}{n} \right) & \text{if } k \geq \theta, \end{cases} \quad (16)$$

When A is rare, N types are in groups with  $n - 1$  other N types, and thus they receive no benefits from the cooperative activities of A types and their marginal fitness is zero.

The threshold punishment game provides a simple model of the payoffs from contingent punishment of non-cooperators. Again individuals interact for  $T$  periods during which they can either cooperate or defect with the same payoffs as in the iterated public goods game. Everyone defects during the first period, and then A types



**Fig. 1.** The minimum number of repetitions necessary for such contingent cooperators to increase when rare as a function of  $r_e$ , the equilibrium relatedness within ethnolinguistic groups. The top two panels give the values for the iterated public goods game, and the bottom two the threshold punishment game. For the latter we assume  $\beta = b$  and  $\chi = c$  and so represent a best case for the evolution of punishment. We also assume that  $\kappa = c$  when  $n = 10$  and  $\kappa = 4c$  when  $n = 40$ . In both cases,  $\theta = 0.25n$ . This means that A types behave cooperatively as long as three others cooperate when  $n = 10$ , and as long as ten others do when  $n = 40$ . The fact that the assumed values of  $\theta$  and  $\kappa$  both scale with  $n$  means that the per capita cost of each act of punishment increases with  $n$ . When groups are small ( $n = 10$ ), a modest number of interactions allows contingent cooperation to increase as long as relatedness is not too low. When  $b/c = 2$  there is no value of  $T$  which allows cooperators with a threshold of  $\theta = 0.25n$  to increase because cooperation doesn't become profitable until  $k = (c/b)n$ , and cooperators lose in the long run when  $k < n/2$ . For the values of  $r_e$  that we are using, the probability mass for these values of  $k$  is high enough that expected payoff is negative for any  $T$ . This doesn't happen for the threshold game because everybody cooperates if the number of punishers exceeds the threshold. The curve for  $b/c = 8$  becomes negative when inclusive fitness alone will allow cooperation to increase. When groups are larger ( $n = 40$ ), a very large number of interactions is necessary when  $r_e$  is low, but a reasonable number when  $r_e$  is higher.



**Fig. 2.** The minimum number of repetitions necessary for such contingent cooperators to increase when rare as a function of  $r_e$ , the equilibrium relatedness within ethnolinguistic groups. The top two panels give the values for the iterated public goods game, and the bottom two the threshold punishment game. For the latter we assume  $\beta = b$  and  $\chi = c$  and so represent a best case for the evolution of punishment. We also assume that  $\kappa = c$  when  $n = 10$  and  $\kappa = 4c$  when  $n = 40$ . In both cases,  $\theta = 0.8n$ . This means that A types behave cooperatively (cooperate in the IPG, punish non-cooperators in the threshold punishment game) as long as eight others cooperate when  $n = 10$ , and as long as 32 others do when  $n = 40$ . When groups are small ( $n = 10$ ) and relatedness is high ( $r_e = 0.1$ ) contingent cooperation and punishment can increase if  $T$  is on the order of 100. For the IPG with  $n = 10$ , the curve for  $b/c = 8$  becomes negative when inclusive fitness alone will allow cooperation to increase. When groups are larger ( $n = 40$ ) contingent cooperation cannot increase unless the number of interactions is very large.

punish defectors at a cost. In general we would expect this cost to increase with group size, but also decrease with the number of punishers. We would also expect economies of scale so that the cost of punishment declines nonlinearly as the number of punishers increases.  $N$  types do not punish. If the number of punishers exceeds a threshold  $\theta$ , all individuals attempt to cooperate on subsequent periods, but occasionally defect due to errors or random individual circumstances that make defection attractive even if punishment is anticipated. A simple model of this situation is

$$V_k^A = \begin{cases} -\kappa & \text{if } k < \theta, \\ -\kappa + (T-1)(\beta - \chi) & \text{if } k \geq \theta, \end{cases} \quad (17)$$

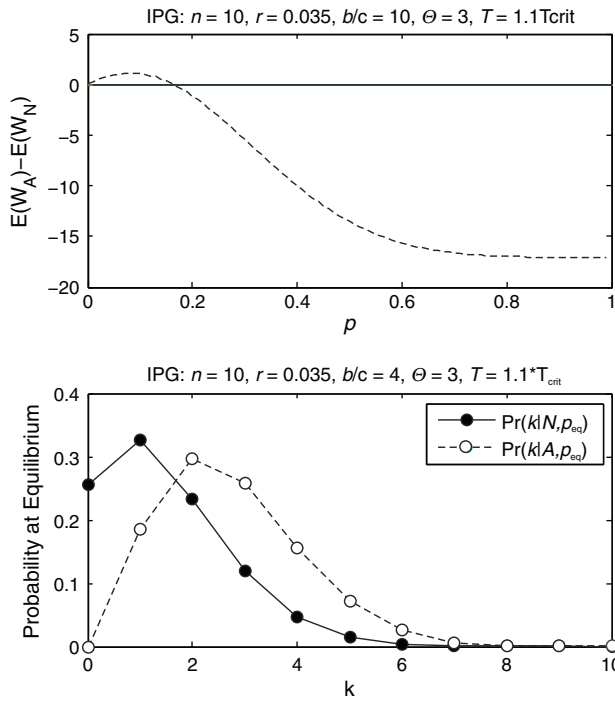
and

$$V_k^N = \begin{cases} 0 & \text{if } k < \theta, \\ (T-1)\beta & \text{if } k \geq \theta, \end{cases} \quad (18)$$

where  $\kappa$  gives the cost of punishing noncooperators and being punished during the first period,  $\beta$  gives the benefits due to induced cooperation and  $\chi$  is the cost of cooperating, monitoring, and punishing occasional defectors. We expect that  $\beta \leq b$  due to occasional defections, and that  $\chi \geq c$  due to the costs of monitoring and enforcement. A more realistic model of contingent punishment in which the effects of group size, economies of scale and errors are made explicit requires an increase in the number of parameters (Boyd et al., 2010). Once again the expected fitness of defectors is zero.

Using (13) to calculate expected fitnesses indicates that the evolution of contingent cooperation is sensitive to the level of

relatedness within ethnolinguistic groups, to band size, and to the threshold necessary for cooperation to persist in a band. Fig. 1 plots the minimum value of  $T$  necessary for A types to increase as a function of this relatedness,  $r_e$  in the iterated public goods game and the threshold punishment game. In both cases  $\theta$ , the minimum number of A types in a band necessary to sustain cooperation is  $0.25n$ . This means that A types cooperate or punish as long as a quarter of their group also does the same. The top two panels plot values for the iterated public goods game and the lower two plot values for the threshold punishment game. The range of  $r_e$  values plotted encompasses a number of empirical estimates of this parameter. The classic Birdsell (1973) model of hunter-gatherer population structure ( $n_e = 500$ ,  $m = 0.05$ ) implies that at equilibrium  $r_e \approx 0.02$ . Birdsell's original calculation was based on little data but is roughly consistent with more recent estimates of group size ( $\approx 1000$  and intermarriage rates  $\approx 0.05$ ) for three central Inuit groups, the Copper Inuit, the Netsilik, and the Iglulik (Damas, 1975). Variable population sizes, unequal sex ratios, deviations from the assumptions of the island model, and a number of other factors might increase these estimates of  $r_e$ . There are two published empirical estimates of  $r_e$  for contemporary human foragers based on genetic data, a low value ( $r \approx 0.01$ ) from Langergraber et al. (2011) and a higher value ( $r \approx 0.07$ ) from Bowles (2006). When  $n = 10$  and  $\theta = 3$ , cooperative strategies increase when rare even when  $r_e$  is fairly low as long as  $T > 100$ , a value that seems reasonable for common interactions like food sharing but less plausible for rare interactions like participation in warfare. For values of  $r_e$  at the higher end of the range, contingent cooperation can increase even when individuals only interact a few times. This should not be surprising because when  $\theta = 3$  cooperation



**Fig. 3.** The top panel plots the difference in expected fitness for A and N types, in units of  $\delta$ , in the iterated public goods game when  $n = 10, r_e = 0.035, b/c = 4, \theta = 3$  and the expected number of interactions is 10% higher than the minimum number of repetitions necessary for such contingent cooperators to invade. The equilibrium frequency,  $p_{eq}$ , of A types occurs when the [id = rv] difference different in expected fitness is zero. The bottom panel shows the probability that there are  $k$  A types in a band conditioned on the focal being either A or N. An A type is more likely to be in a band with enough A types to sustain cooperation, and this just compensates for the fact that A types suffer more than N types when cooperation is not sustained.

(or punishment) is, in essence, a three person interaction. When  $n$  is larger, the evolution of contingent cooperation is extraordinarily sensitive to the relatedness within ethnolinguistic groups. Contingent cooperation cannot increase when  $r_e$  is at the low end of the range, but increases under plausible conditions at the high end of the range.

Less tolerant contingent strategies increase only when relatedness within ethnolinguistic groups is high and bands are small. Fig. 2 plots

the minimum value of  $T$  necessary for A types to increase as a function of this relatedness, but now assuming that  $\theta$ , the minimum number of A types in a band necessary to sustain cooperation, is  $0.8n$ . This means that A type cooperates or punishes as long as 80% of their group also does the same. When bands are small ( $n = 10$ ) cooperation can increase if  $r_e$  is at the upper end of the range and individuals interact many times. Otherwise there are no plausible conditions which allow A types to increase.

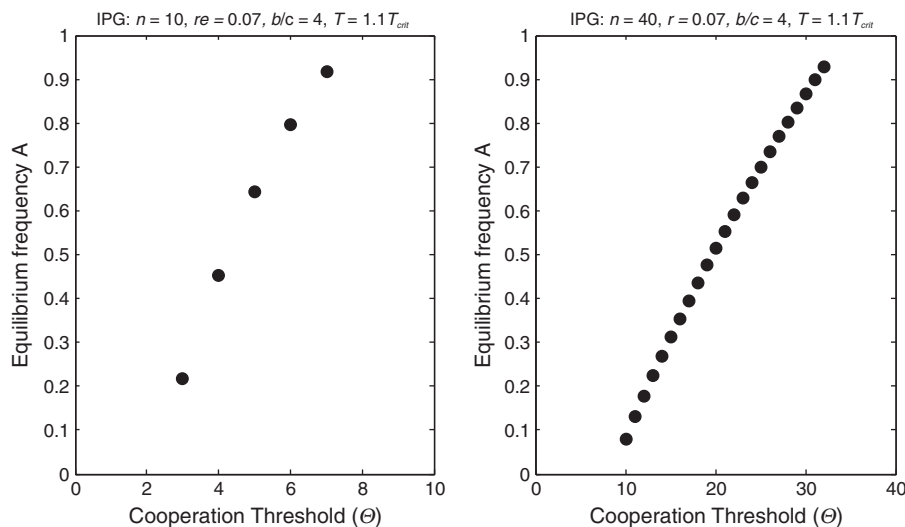
### 3. Equilibrium frequencies of cooperation

Even when they can invade when rare, contingent cooperators typically do not eliminate non-cooperative types. Instead the population usually evolves to a stable equilibrium at which both types coexist and have the same expected fitness. When A types are rare, N types virtually always find themselves in bands with  $n - 1$  other N types, and so reap virtually no benefit from cooperative behavior. In contrast, A types are occasionally in groups with a sufficient number of A types to sustain cooperation, and as long as the expected fitness benefits from such groups compensates for the losses in bands that do not sustain cooperation, A types can increase. As they do, however, more and more N types find themselves in bands in which cooperation is sustained. N types are less likely to be in such bands than A types, but they also don't pay the cost of cooperation (or punishment). Eventually, these two fitness components balance, and the population reaches a stable mixed equilibrium. See Fig. 3 for an illustration.

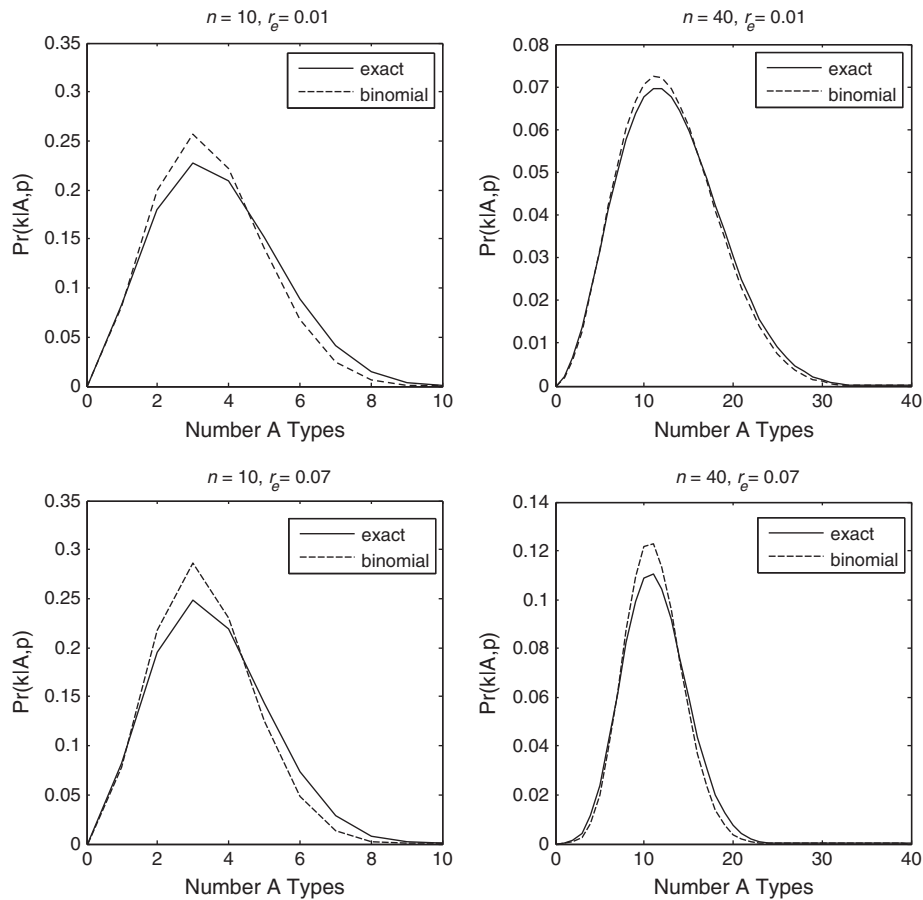
The equilibrium frequency of contingent cooperators depends on the threshold number required for cooperation (Fig. 4). When the threshold is low, so too is the frequency of cooperators at equilibrium; when the threshold is high, the frequency of cooperators at equilibrium is high. This means that the strategies that can invade at lower levels of relatedness within ethnolinguistic groups reach a low frequency at equilibrium. Low equilibrium frequencies of cooperation mean that a substantial fraction of bands do not sustain cooperation because there are no sufficient numbers of A types.

### 4. Robustness of the model

The model analyzed in this paper incorporates a number of assumptions that make it possible to derive a simple recursion for the distribution of types across bands. We think that the three most



**Fig. 4.** Equilibrium frequency of A types in the IPG plotted against  $\theta$ , the number of A types necessary for cooperation or punishment to be sustained. The number of repetitions,  $T$ , is set at 10% above the minimum number for A types to increase,  $r_e = 0.07$ , and  $b/c = 4$ . In the left panel  $n = 10$  and in the right  $n = 40$ . As explained in Fig. 1, when the threshold is small invasion is not possible no matter how large  $T$  is.



**Fig. 5.** The probability that there are  $k$  A types in a group given that the focal is type A and the frequency of A types in the population is  $p$ ,  $\Pr(k | A, p)$  assuming that bands exchange migrants at a rate  $m_b$ , compared with the model assuming random band formation (“binomial”). All panels assume  $p = 0.25$  and  $m_b = 0.5$ . In the top two panels  $r_e = 0.01$  and in the bottom two panels  $r_e = 0.07$ . In the left two panels  $n = 10$  while in the right two  $n = 40$ .

important are: (1) random formation of bands, (2) weak selection, and (3) island model migration among ethnolinguistic groups. In this section we consider whether violations of these simplifying assumptions are likely to have a large effect on the results.

#### 4.1. Random formation of bands

While foragers often move from one band to another during their lives, it is unlikely that bands are random samples from the ethnolinguistic group to which they belong. Instead, people live with a mix of relatives and non-relatives, and this will lead to greater relatedness within bands. To test the sensitivity of the model to non-random band formation, we constructed a two level model in which every generation bands exchange a fraction  $m_b$  of migrants with other bands in their ethnolinguistic group. Ethnolinguistic groups are linked in an island model as before, and the distribution of frequencies of the two types across ethnolinguistic groups is well approximated by a beta distribution. However, because bands are small and migration rates between bands are substantial, the beta distribution does not provide a good approximation for the probability distribution of the number of A types across bands. Instead we used the matrix methods described in Wakeley (2003) eqn 4 to calculate the exact distribution of the number of A types in bands given the type of a focal and the frequency  $x$  in the ethnolinguistic group,  $\nu(k|A, x) = k\nu(k|x)/(nx)$  and  $\nu(k|N, x) = (n - k)\nu(k|x)/(n(1 - x))$ , where  $\nu(k|x)$  is the probability that a band chosen at random in an ethnolinguistic group in which A types have frequency  $x$  has  $k$  A

types. Then the probability that a type A focal is in a band with  $k$  A types is

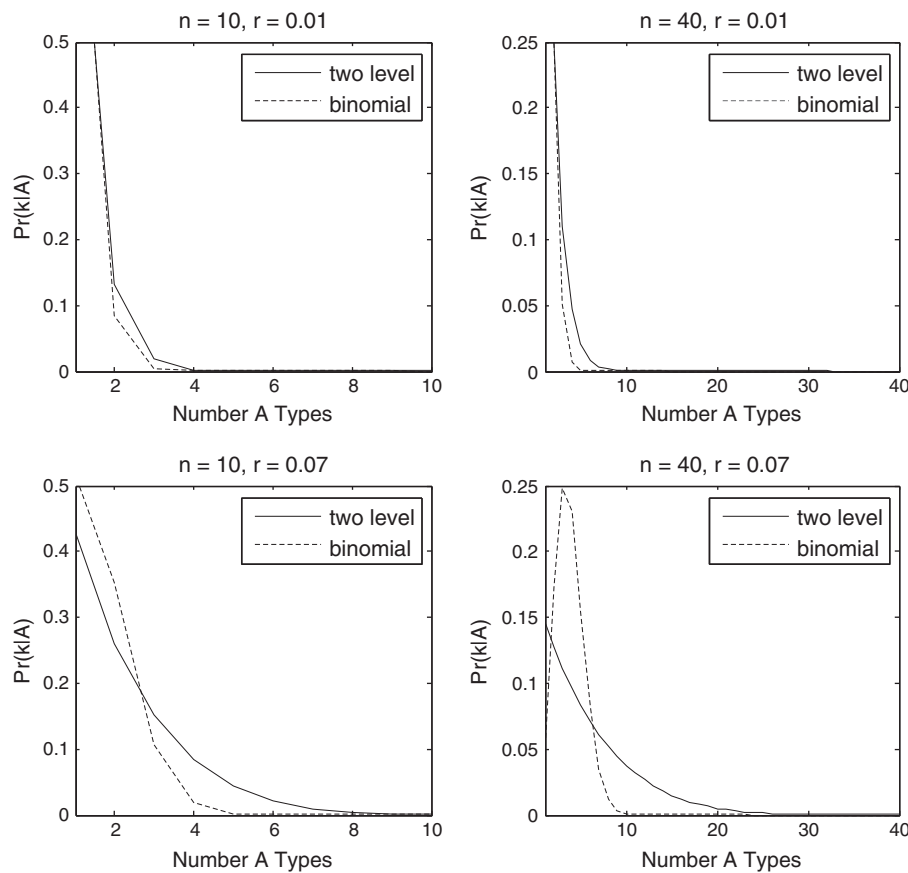
$$\Pr(k|A, p) = \int_0^1 \nu(k|A, x) \text{beta}(x|p + 1, q) dx. \tag{19}$$

The integral on the right hand side of (19) can be evaluated numerically and compared to the simpler model of band formation used above. As is illustrated in Fig. 5, when the migration rate between bands is greater than 0.5 the distributions assuming random band formation and limited between band migration are very similar, even when  $r_e$  is small. Given that the fraction of first degree relatives in bands is typically less than 10% (Hill et al., 2011) this result suggests that the random band model provides an adequate approximation to the more realistic model.

#### 4.2. Weak selection

The model analyzed above assumes that selection is sufficiently weak that the time scale on which it changes gene frequencies is much longer than the time scale on which migration and drift adjust the distribution of types across ethnolinguistic units. Put another way, selection responds to the long run average distribution of behavioral strategies across the entire population. The data suggest that migration rates among ethnolinguistic groups are fairly small, and this means that the model applies only if selection is fairly weak.

It is difficult to analyze models with strong selection, and so it is not clear what is the effect of this assumption. However, we conjecture that stronger selection will lead to less contingent



**Fig. 6.** The probability that there are  $k$  A types in a group given that the focal is type A and that A types are rare  $\Pr(k|A)$  assuming the two level population structure studied in the paper compared with a binomial distribution assuming a fixed value of  $r_e$ . In the top two panels  $r_e = 0.01$  while in the bottom two  $r_e = 0.07$ . In left two panels  $n = 10$  while in the right two panels  $n = 40$ . When  $r_e$  is small the two distributions are similar, but when  $r_e$  is larger the binomial assumption underestimates the extent to which contingent cooperation can increase when rare.

cooperation. Suppose selection is strong enough that the time scale of selection is shorter than that of drift and migration. This means that ethnolinguistic units are demes, and genes will compete only with genes in their ethnolinguistic units. Since bands are formed at random, there is no assortment within ethnolinguistic units, and therefore contingently cooperative strategies that are altruistic when rare will not increase in frequency.

#### 4.3. Island model migration among ethnolinguistic units

Ethnolinguistic groups are large and may often occupy sizable territories. This suggests that the island model in which ethnolinguistic groups exchange migrants with all other ethnolinguistic groups in the population is unlikely to be correct, and a model with local migration would be more accurate. Such a model would be difficult to analyze analytically. Remember, however, that our estimates of  $r_e$  do not depend on the assumption of island model migration; they are empirical. It is only the beta distribution that depends on the assumption of island model migration. So as long as a model with limited migration generates a distribution with the same general shape as the beta, we conjecture that the island model will be an adequate approximation.

## 5. Discussion

In this paper we have derived a simple recursive set of formulas, (9), (11) and (12) (simplified to (13), when one type is rare) that gives the distribution of genotypes across bands in a two level model that captures important aspects of population structure observed in

Holocene hunter-gatherers. It is easy to use this distribution to calculate expected fitness of strategies that support contingent cooperation within hunter-gatherer bands.

Assuming that this two level population structure also characterized Pleistocene foragers, these results have a number of implications for the evolution of contingent cooperation in human societies:

1. Using the two level model predicts that the strategies that cooperate contingent on past cooperation by a minority can increase when rare only if groups are quite small and relatedness is fairly high ( $r > 0.05$ ). The expected payoff of contingent cooperators is very sensitive to the relatedness within ethnolinguistic groups. When this is on the order of 0.01 as estimated by Langergraber et al. (2011), cooperative strategies can increase only if the bands are quite small and long run payoffs very substantial. On the other hand, when this is larger as estimated by Bowles (2006) contingent cooperators can increase under a wider range of conditions, although these conditions are still very sensitive to group size.
2. Strategies that cooperate contingent on past cooperation by a minority of individuals can increase under plausible conditions. Strategies that require the cooperation of a majority cannot.
3. If cooperative strategies can invade and relatedness within ethnolinguistic groups is not too high, the population evolves to a stable equilibrium at which both types are present. The frequency of cooperative types at such equilibria depends on their degree of tolerance. Highly tolerant strategies that cooperate even if only a small number of others cooperated in the past occur at low frequency at equilibrium, while less tolerant strategies can achieve high frequency.



4. Since intolerant strategies cannot invade under plausible conditions, this result suggests that wide spread cooperation in sizable groups is unlikely to be supported by contingent cooperation. However, if a minority of punishers are sufficient to motivate all band members to cooperate, contingent punishment strategies that can invade under plausible conditions can also support widespread cooperation.
5. Existing work (Boyd & Richerson, 1988; Boyd et al., 2010) overestimates the amount of relatedness necessary for contingent cooperation to increase when rare. In these models, band members are assumed to be sampled from within a population with a fixed level of relatedness. Because this assumption ignores variation between ethnolinguistic groups it also underestimates the strength of selection that favors cooperative strategies which are altruistic when rare. The magnitude of the underestimate depends strongly on the relatedness within ethnolinguistic groups and therefore the amount of variation among ethnolinguistic groups. When this is small (i.e. relatedness within ethnolinguistic groups is low), there is little qualitative difference, but when it is large (relatedness within ethnolinguistic groups is higher), previous work substantially underestimates the probability that groups with a significant number of cooperators will be formed (Fig. 6).

## References

- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211, 1390–1396.
- Berndt, R. M. (1959). The concept of “the tribe” in the Western Desert of Australia. *Oceania*, 30, 81–107.
- Birdsell, J. B. (1973). A basic demographic unit current anthropology, 14, 337–356.
- Bowles, S. (2006). Group competition, reproductive leveling, and the evolution of human altruism. *Science*, 314, 1569–1572.
- Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, 328, 617–620.
- Boyd, R., & Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, 132, 337–356.
- Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13, 171–195.
- Clutton-Brock, T. (2009). Cooperation among non-kin in animal societies. *Nature*, 462, 51–57.
- Crow, J. F., & Kimura, M. (1970). An introduction to populations genetics theory. New York: Harper & Row.
- Damas, D. (1975). Demographic aspects of central Eskimo marriage practices. *American Ethnologist*, 2, 409–418.
- Doebeli, M., & Hauert, C. (2005). Models of cooperation based on the prisoner's dilemma and the snowdrift game. *Ecology Letters*, 8, 748–766.
- Gardner, A., & West, S. A. (2006). Demography, altruism and the benefits of budding. *Journal of Evolutionary Biology*, 19, 1707–1716.
- Gardner, A., West, S. A., & Wild, G. (2011). The genetical theory of kin selection. *Journal of Evolutionary Biology*, 24, 1020–1043.
- Goodnight, K. (1992). The effect of stochastic variation on kinselection in a budding-viscous population. *American Naturalist*, 140, 1028–1040.
- Hart, D., & Clark, A. (2007). Principles of population genetics. Sunderland, MA: Sinauer.
- Hill, K. R., Walker, R. S., Bozicevic, M., Eder, J., Headland, T., Hewlett, B., et al. (2011). Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science*, 331, 1286–1289.
- Joshi, N. V. (1987). Evolution of cooperation by reciprocation within structured demes. *Journal of Genetics*, 66, 69–84.
- Langergraber, K., Schubert, G., Rowney, C., Wrangham, R., Zommers, S., & Vigilant, L. (2011). Genetic differentiation and the evolution of cooperation in chimpanzees and humans. *Proceedings of the Royal Society B*, 278, 2546–2552.
- Mathew, S., & Boyd, R. (2011). Punishment sustains large-scale cooperation in prestate warfare. *Proc. Nat. Acad. Sci. USA*, 108, 11375–11380.
- Mathew, S., Boyd, R., & vanVeelen, M. (2013). Human cooperation among kin and close associates may require enforcement of norms by third parties. In P. J. Richerson, & M. Christiansen (Eds.), *Cultural Evolution* (pp. 45–60). Cambridge, MA: MIT Press Strngmann Forum Report 12, series editor J. Lupp.
- McElreath, R., & Boyd, R. (2007). *Mathematical models of social evolution: a guide for the perplexed*. Chicago: University of Chicago Press.
- Nuligak, M. (1971). *I, Nuligak*. NY: Simon Schuster.
- Schonmann, R. H., Vicente, R., & Caticha, N. (2013). Altruism can proliferate through population viscosity despite high random gene flow. *PLoSOne*, 8, 1–10.
- Schonmann, R.H., Boyd, R and Vicente, R. (in prep). The evolution of cooperation under local regulation and non-additive gene action: an extension of Hamilton's fundamental ideas.
- Swezey, S. L., & Heizer, R. F. (1977). Ritual management of salmonid fish resources in California. *Journal of California Archaeology*, 4, 6–29.
- Taylor, P. D. (1992). Altruism in viscous population – an inclusive fitness approach. *Evolutionary Ecology*, 6, 352–356.
- Tindale, N. B. (1974). *Aboriginal tribes of Australia: their terrain, environmental controls, distribution, limits, and proper names*. Canberra: Australian National University Press.
- van Veelen, M. (2009). Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. *Journal of Theoretical Biology*, 259, 589–600.
- Wakeley, J. (2003). Polymorphism and divergence for island-model species. *Genetics*, 163, 411–420.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16, 97–159.