Ⓔ

Cultural Transmission Can Inhibit the Evolution
of Altruistic Helping

Laurent Lehmann,[1],[*] Marcus W. Feldman,[1],[†] and Kevin R. Foster[2],[‡]

1. Department of Biological Sciences, Stanford University,
Stanford, California 94305;
2. Bauer Center for Genomics Research, Harvard University,
Cambridge, Massachusetts 02138

ABSTRACT: The study of culturally inherited traits has led to the suggestion that the evolution of helping behaviors is more likely with cultural transmission than without. Here we evaluate this idea through a comparative analysis of selection on helping under both genetic and cultural inheritance. We develop two simple models for the evolution of helping through cultural group selection: one in which selection on the trait depends solely on Darwinian fitness effects and one in which selection is driven by nonreproductive factors, specifically imitation of strategies achieving higher payoffs. We show that when cultural variants affect Darwinian fitness, the selection pressure on helping can be markedly increased relative to that under genetic transmission. By contrast, when variants are driven by nonreproductive factors, the selection pressure on helping may be reduced relative to that under genetic inheritance. This occurs because, unlike biological offspring, the spread of cultural variants from one group to another through imitation does not reduce the number of these variants in the source group. As a consequence, there is increased within-group competition associated with traits increasing group productivity, which reduces the benefits of helping. In these cases, selection for harming behavior (decreasing the payoff to neighbors) may occur rather than selection for helping.

Keywords: altruism, helping, harming, cultural relatedness, cultural transmission, spatially subdivided populations.

* E-mail: lehmann@stanford.edu.

† E-mail: marc@charles.stanford.edu.

‡ E-mail: kfoster@cgr.harvard.edu.

Cultural group selection, whereby differential group success results from the expression of different cultural traits, has been proposed as a major force driving the evolution of "altruistic" helping in humans (Fehr et al. 2002; Henrich 2004; Richerson and Boyd 2005). While genetically inherited traits tend to follow strict vertical transmission from parent to offspring, culturally inherited traits may be transmitted obliquely from adults outside the nuclear family to individuals of the offspring generation. Cultural variants may also be transmitted horizontally between individuals of the same generation. For instance, birds learn song dialects from neighboring individuals (Bonner 1980), chimpanzees imitate group mates that have the skills of "termiting" with tools (Goodall 1966), and humans gain knowledge from unrelated individuals living in present and past generations (Cavalli-Sforza and Feldman 1981; Lumsden and Wilson 1981; Boyd and Richerson 1985). A direct consequence of these nonvertical transmission schemes is that they can rapidly homogenize behavior within groups, and as a result, most individuals in a group may display the same cultural trait (i.e., high "cultural inbreeding" sensu Cavalli-Sforza and Feldman 1981 [p. 352] or "cultural relatedness" sensu Allison 1991). For instance, if offspring learn a cultural trait by copying a few individuals in their group (the "teachers" or "leaders") instead of each offspring independently copying its parents, the within-group variance of cultural traits decreases and the between-group variance increases (see also Cavalli-Sforza and Feldman 1973). Thus, when helper and defector traits are culturally inherited, helpers will tend to occur together in a group, and so will defectors or cheaters. The resulting within-group uniformity reduces the tendency of cheaters to exploit helpers and concomitantly increases the evolutionary success of helpers. This argument has led to the view that cultural group selection augments any effects of genetic transmission in favoring the evolution of "altruism" among humans, especially in groups of large size where genetic group selection is inefficient (e.g., Fehr et al. 2002; Henrich 2004; Boyd and Richerson 2005; Richerson and Boyd 2005).

In summarizing a major section of their book, Boyd

and Richerson (2005, p. 17) say, "In part III we discuss how conformity reduces within-group cultural variation, making group-level selection a more plausible process than group selection on genes is usually thought to be." But is it generally true that group-level selection is more plausible through cultural processes than genetic processes? This would be a decisive finding because it would suggest that cultural transmission should be directly associated with an increased tendency to help others (i.e., culture begets cooperation). In order to confirm this conjecture, we require a direct comparison of social evolution under the two transmission schemes, genetic and cultural. This comparison has been made for panmictic populations by Feldman et al. (1985), who showed that in fact it may be more difficult for helping to increase in frequency under cultural transmission than under genetic assumptions. A preliminary analysis suggested further that the same may be true for geographically subdivided populations (Lehmann et al. 2007). In this article, we build on these previous studies and compare genetic and cultural transmission in geographically subdivided populations by developing simple models of cultural inheritance of helping traits that are easily compared with classical models of genetically inherited traits. This enables us to assess directly whether cultural transmission of helping will result in stronger or weaker selective pressure on helping than occurs under genetic transmission.

The directional (or systematic) change in frequency of a cultural variant (as opposed to strict cultural drift for a neutral or unbiased cultural variant; e.g., Cavalli-Sforza and Feldman 1981) can be affected by a multitude of factors, many of which do not affect genetic traits. At the risk of over simplification, we divide these factors into two non–mutually exclusive categories. First, systematic change in the frequency of cultural variants can be caused by their effects on Darwinian fitness, as is the case with biological traits (effects on reproduction; e.g., Cavalli-Sforza and Feldman 1981; Boyd and Richerson 1985). Most simply, the behavior of having many offspring can spread and lead to an increase in the frequency of some religious practices in a population where the practices are carried out by individuals with many offspring. For example, one study found that Australian Roman Catholics had a 40% higher reproductive fitness than nonreligious Australians (Kirk et al. 2001, their table 2). If differential reproduction were the only cause of systematic change in the frequency of cultural variants, cultural evolution would mainly be driven by vertical transmission from parent to offspring, and any unbiased horizontal or oblique transmission (variants are assumed neutral during this stage) would only alter the distribution of traits within and between groups as a result of random sampling (i.e., cultural drift).

The second category of processes that cause directional change in the frequencies of cultural traits are those that occur independently of reproductive effects. These processes operate through some other form of differential variant success, such as a cognitive preference or biased cultural transmission (e.g., conformist or other frequency-dependent transmission or assimilation functions of cultural variants; see Lumsden and Wilson 1980). Such nonreproductive differential success can change cultural trait frequencies through vertical, horizontal, or oblique transmission. For instance, individuals might have a tendency to imitate/copy cultural traits that are not predictive of reproductive performance. Such imitation is often assumed to occur by payoff comparison, whereby individuals copy the strategies of those individuals around them that are perceived to have higher payoffs (Boyd and Richerson 1985; Rogers 1995; Hofbauer and Sigmund 1998; Kendal et al. 2006). This could be, for instance, imitating the successful or the wealthiest, where a social helping trait might be a timely gift of money to others that are particularly in need at that moment. Under these situations, a helping trait may increase in frequency without necessarily affecting Darwinian fitness.

We develop two simple models for the cultural evolution of helping behaviors, whereby an individual may increase the payoff to other individuals in some way. The term "cooperation" is sometimes used in a broad sense, interchangeably with helping (e.g., Boyd and Richerson 2005; West et al. 2007), and the helping behaviors we study here are primarily altruistic sensu Hamilton (1964, 1970) with respect to the currency that drives selection. In the first model, the directional change in the frequency of helping is due solely to its effect on reproductive fitness, while in the second, the changes are caused by nonreproductive differential success. Our goal is to represent two extreme cases in a continuum of possible modes of cultural transmission and to compare them with genetic transmission. We find some confirmation of the view that, relative to genetically based group selection, cultural group selection may increase the intensity of selection for helping behaviors. When cultural traits affect only reproductive success, oblique or horizontal cultural transmission can markedly reduce within-group variation, thereby increasing the selective pressure on helping relative to that under simple genetic selection. We find a starkly different picture for the evolution of cultural traits with nonreproductive differential success. When transmission is based on a tendency to imitate or learn from individuals with higher payoffs, an assumption that is often made for social models (Ellison 1993; Kandori et al. 1993; Hofbauer and Sigmund 1998; Boyd and Richerson 2002; Boyd et al. 2003; Hauert and Doebeli 2004; Kendal et al. 2006), selection for helping may be markedly reduced relative to that under genetic

transmission. In fact, harming behavior between members of a group (as opposed to helping) may be selected for under this transmission mode.

Model

Life Cycle

Suppose individuals live in a population consisting of an infinite number of groups, each of finite size N. We consider a single haploid cultural "locus" controlling the expression of a helping behavior that is inherited culturally. We assume that two cultural "variants" segregate at this cultural locus; that is, we consider a dichotomous trait (e.g., Cavalli-Sforza and Feldman 1981; Boyd and Richerson 1985). An individual bearing the "helping" variant (e.g., A) expresses a behavior that results in a payoff cost C to himself/herself but generates a payoff benefit B that is shared equally among all group members except the actor. An individual bearing the "defector" variant (e.g., a) takes the payoff benefit but pays no cost. These assumptions greatly simplify the complexities of the expression of cultural traits, their inheritance, and the evaluation of how differential success translates into fitness costs and benefits. They do, however, provide a useful approximation that allows us to disentangle the effects of diverse demographic regimes and cultural transmission rules on the evolution of the trait. It also appears that the conclusions from such simple discrete models retain qualitative relevance in more sophisticated cultural scenarios, for example, continuous traits and nondiscrete mental representations of traits (Cavalli-Sforza and Feldman 1981; Henrich and Boyd 2002). We also assume that individuals, regardless of the cultural variants they carry, have a high fecundity. This assumption is obviously violated in many human societies; it is adopted here to avoid the intricacies of demographic stochasticity (fluctuation of group size), which would make qualitative comparisons between the models much more difficult.

The events of our idealized life cycle occur in the following order. (1) Reproduction occurs, with each adult individual producing a large number of juveniles according to its phenotype and the phenotypes of its group mates. Juveniles inherit/assimilate the cultural variant of their parent (i.e., vertical transmission of cultural variants). (2) Juveniles disperse independently of each other with probability m to another group. (3) Density-dependent regulation occurs, and exactly N juveniles reach adulthood in each group. (4) Horizontal transmission of cultural variants occurs among the adult individuals in the population.

One can imagine many different horizontal transmission modes occurring during the last stage of the life cycle. Here, we will consider only two special cases, both of which have been studied previously. (i) The cultural variants affect only the reproductive fitness of the actor and recipients. In this case, horizontal transmission can occur (one individual might copy/imitate another), but the tendency to copy is not affected by whether the "template" or "exemplar" individual is a helper or a defector (C and/or B do not affect cultural transmission); hence, the cultural variants are neutral during horizontal transmission (i.e., there is no systematic change of variant frequency at the population level). (ii) The phenotypic effects of cultural variants (C and B) affect their probability of horizontal transmission but have no effects on reproduction. For this case, we analyze two different payoff-based imitation rules during horizontal transmission, whereby individuals tend to adopt cultural variants by comparing the payoffs of different classes of individuals: a "group payoff comparison" rule, where individuals compare their payoffs with the average payoff in a group, and a "pairwise payoffs comparison" rule, where individuals compare their payoff with the payoff of a randomly encountered individual. These two updating schemes will be described in detail during the mathematical analysis.

To aid in the interpretation of our results, we first recall that in the absence of horizontal transmission and with only vertical transmission (which is equivalent here to haploid genetic inheritance), the condition for the invasion of the helping variant A under the life cycle described above is given under weak selection by

$$-C > 0. \tag{1}$$

Helping is selected for only if the actor's fecundity—that is, the number of its juveniles counted before the density-dependent regulation stage—is increased (Taylor 1992a, 1992b). Taylor's $-C > 0$ rule is independent of the spatial structure of the population (see also Wilson et al. 1992). It can be used as a yardstick with which the effect of changing assumptions on the evolution of helping behaviors in populations subject to limited dispersal can be compared. For instance, costly helping can evolve when it results in group size expansion by reducing group extinction rates (i.e., "selective extinction"; e.g., Eshel 1972; Aoki 1982; Bowles 2006) or by increasing individual productivity through the effect of the helping behavior on various demographical variables (e.g., van Baalen and Rand 1998; Le Galliard et al. 2003; Lehmann et al. 2006). Selection on helping is also increased by changing assumptions about individual life history (such as introducing overlapping generations, non-Poisson progeny distribution, sex-biased dispersal), the mode of dispersal, the mode of competition between groups, and the mode of group fission (e.g., Rogers 1990; Taylor and Irwin 2000; Irwin and Taylor 2001; Gardner and West 2006; Ohtsuki et al. 2006; Grafen 2007;

Lehmann and Balloux 2007; van Veelen and Hopfensitz 2007). Recognition of kin or punishment of defectors within groups (e.g., Boyd et al. 2003; Axelrod et al. 2004; Jansen and van Baalen 2006; Rousset and Roze 2007) and niche construction effects (e.g., Lehmann 2006; Wakano 2007) also favor the evolution of costly helping. Here, we ask whether the two different modes of horizontal transmission introduced above will result in more relaxed or more stringent conditions for the evolution of helping than predicted by the $-C > 0$ rule.

Price Equation

The average frequency p of variant A in the population after one iteration of the life cycle is given by

$$p' = \mathrm{E}[w_{ij}p^*_{ij}], \qquad (2)$$

where the prime denotes the value in the next generation, the expectation operator $\mathrm{E}[\cdot]$ refers to an average over all groups in the population and all individuals within groups ($p = \mathrm{E}[p_{ij}]$), w_{ij} is the fitness function giving the expected number of successful offspring of individual i of group j, and p^*_{ij} is the frequency (0 or 1) of variant A in that individual before reproduction but after horizontal cultural transmission. Here, $p^*_{ij} = p_{ij} + \Delta p_{ij}$, where p_{ij} is the frequency (0 or 1) of variant A in individual ij before cultural transmission and Δp_{ij} is the change in the number of A's (0, 1, or -1) in individual i of group j as a result of horizontal cultural transmission. Substituting p^*_{ij} into equation (2), we obtain $p' = \mathrm{E}[w_{ij}p_{ij}] + \mathrm{E}[w_{ij}\Delta p_{ij}]$. Noting that $\mathrm{E}[w_{ij}\Delta p_{ij}] = \mathrm{E}[w_{ij}]\mathrm{E}[\Delta p_{ij}]$ because horizontal transmission occurs before reproduction, so that Δp_{ij} is independent of the fitness of individual ij, and that the expectation of the fitness function over all individuals is equal to 1 because population size is constant ($\mathrm{E}[w_{ij}] = 1$), we obtain the change in frequency of variant A:

$$\Delta p = \mathrm{Cov}(w_{ij}, p_{ij}) + \mathrm{E}[\Delta p_{ij}], \qquad (3)$$

which is a special case of the Price equation (Price 1970). This equation emphasizes that the per generation change in the frequency of A depends on two terms: first, on the covariance $\mathrm{Cov}(w_{ij}, p_{ij})$ between the fitness of an individual and the frequency of A in that individual, and second, on the expected change $\mathrm{E}[\Delta p_{ij}]$ in variant A frequency as a result of horizontal cultural transmission (several rounds of horizontal cultural transmission could actually occur within a generation).

Under alternative (i) of our life cycle, there is no systematic change of variant frequency through horizontal transmission ($\mathrm{E}[\Delta p_{ij}] = 0$), and the change in variant A frequency is entirely accounted for by the covariance between fitness and variant frequency before cultural transmission, namely

$$\Delta p = \mathrm{Cov}(w_{ij}, p_{ij}). \qquad (4)$$

By contrast, under alternative (ii) of our life cycle, variant A does not affect fitness ($\mathrm{Cov}(w_{ij}, p_{ij}) = 0$), and its systematic change in frequency is entirely accounted for by cultural transmission ($\Delta p = \mathrm{E}[\Delta p_{ij}] = \mathrm{E}[p^*_{ij} - p_{ij}]$). Let $p^*_{ij} = t_{ij}p_{ij}$, where t_{ij} denotes the transmission function giving the expected number of individuals that have inherited the cultural variant from individual i in group j; noting that $\mathrm{E}[t_{ij}] = 1$ because population size is constant, we can write

$$\Delta p = \mathrm{Cov}(t_{ij}, p_{ij}). \qquad (5)$$

In order to evaluate the fitness function w_{ij} and the transmission function t_{ij}, we need a payoff function for the individuals in the population. The payoff to individual i in group j (relative to the payoff to an individual that neither gave nor received helping) will be written as

$$1 + f_{ij} = 1 - Cp_{ij} + \frac{B}{N-1}\sum_{k \neq i} p_{kj}. \qquad (6)$$

That is, individual ij pays a cost C if he/she carries variant A and receives a fraction $1/(N-1)$ of the benefit B generated by each individual (excluding himself) that helps in group j. For the "Darwinian" model, $1 + f_{ij}$ is simply the fecundity of individual ij, while for the "imitation" model, it represents a preference that is ascribed to that individual. We will also use the payoff functions $1 + f_j$, namely the average of $1 + f_{ij}$ in group j (see eq. [A10] in the appendix in the online edition of the *American Naturalist*). Finally, we will need a payoff function that averages $1 + f_j$ over all groups excluding group j from the average. Because of our assumption that there is an infinite number of groups in the population, an average over all groups, excluding a single focal group, converges to the average payoff $1 + f$ over all groups in the population (see eq. [A11]). For notational simplicity, we will thus use $1 + f$ for any average over all groups excluding a single focal group.

Helping Affects Reproduction

Following previous work (e.g., Frank 1998; Rousset 2004), we now derive the direct fitness function w_{ij} of individual ij. This depends on both the expected number of its offspring reaching adulthood in group j and on those reaching adulthood in other groups after dispersing. A number $(1 - m)(1 + f_{ij})$ of the offspring of individual i from group

j remain in group j and compete against an average number $(1 - m)(1 + f_j)$ of offspring produced in that group and an average number $m(1 + f)$ of immigrant offspring produced in the other groups. A number $m(1 + f_{ij})$ of the offspring of individual ij disperse and compete against an average number $1 + f$ of offspring. Collecting all terms, the fitness of individual ij is

$$w_{ij} = \frac{(1 - m)(1 + f_{ij})}{(1 - m)(1 + f_j) + m(1 + f)} + \frac{m(1 + f_{ij})}{1 + f}. \quad (7)$$

In order to simplify our analysis, we now assume that the phenotypic effects C and B on fecundity are of small order δ. This "weak selection" assumption (e.g., Hamilton 1964; Bürger 2000; Kirkpatrick et al. 2002; Roze and Rousset 2003) implies that f_{ij}, f_j, and f will be of order δ. Recalling that for small x we have $y/(1 + x) \simeq y(1 - x)$, we approximate (by neglecting second- and higher-order terms) the increments and/or decrements in fitness of individual ij resulting from the expression of the cultural variant A by all individuals in the population carrying it as

$$w_{ij} \simeq 1 + (f_{ij} - f) - (1 - m)^2(f_j - f). \quad (8)$$

Helping Affects Cultural Preferences

We now derive the direct transmission function t_{ij} of individual i in group j, which depends on the expected number of individuals that adopt its cultural variant through imitation. We express the transmission function as

$$t_{ij} = \sum_{l, k} \Pr(kl \leftarrow ij), \quad (9)$$

where $\Pr(kl \leftarrow ij)$ is the probability that individual k from group l adopts (copies) the cultural variant of individual i from group j, and the sum runs over all groups in the population and all individuals within groups. Hence, t_{ij} represents the total probability that the cultural variant carried by individual ij is copied.

As mentioned in "Life Cycle," we will investigate the effects of two different payoff-based imitation rules on the probabilities of horizontal transmission. Under the "group payoff comparison" rule, with probability $1 - m_{ct}$, a focal individual compares the payoff of a randomly chosen member of his/her group (including himself) with that of the average payoff in his/her group and adopts the variant of the chosen member with probability given by the payoff to that individual relative to the average payoff in the focal individual's group. From these assumptions, the probability $\Pr(kj \leftarrow ij)$ that any individual k from group j adopts

the cultural variant of individual i from group j is given by

$$\Pr(kj \leftarrow ij) = \frac{(1 - m_{ct})(1 + f_{ij})}{N(1 + f_j)}, \quad (10)$$

because with probability $(1 - m_{ct})/N$, individual ij is chosen for payoff comparison by individual kj, in which case the cultural variant of ij is copied with probability $(1 + f_{ij})/(1 + f_j)$. With complementary probability m_{ct}, a focal individual compares the payoff of a randomly chosen individual from another group (nonfocal group) with that of the average payoff over all different groups and adopts the variant of the chosen individual with a probability given by the relative payoff to that individual. Therefore, the probability that individual k from group h (where $h \neq j$) adopts the cultural variant of individual i from group j is

$$\Pr(kh \leftarrow ij) = \frac{m_{ct}(1 + f_{ij})}{(n_d - 1)N(1 + f)}, \quad (11)$$

where $m_{ct}/(n_d - 1)N$ is the probability that individual ij is chosen for payoff comparison by individual kh ($h \neq j$) and where n_d is the (very large) number of groups in the population (i.e., $n_d \to \infty$). Substitution of equations (10) and (11) into equation (9) gives

$$t_{ij} = \frac{(1 - m_{ct})(1 + f_{ij})}{1 + f_j} + \frac{m_{ct}(1 + f_{ij})}{1 + f}. \quad (12)$$

Under complete migration ($m = 1$) and with complete cultural mixing ($m_{ct} = 1$), the fitness function (eq. [7]) and the transmission function (eq. [12]) become identical: $w_{ij} = t_{ij} = (1 + f_{ij})/(1 + f)$. That these two models become similar in a panmictic population is a standard result of population genetics (Hofbauer and Sigmund 1998; Kendal et al. 2006). However, as soon as the population is structured (i.e., $m < 1$ and $m_{ct} < 1$), the two models diverge. Intuition as to the difference between them derives from the assumption that f_{ij}, f_j, and f are of small order δ, and following the same argument that led to equation (8), we have from equation (12) the weak selection approximation

$$t_{ij} \simeq 1 + (f_{ij} - f) - (1 - m_{ct})(f_j - f). \quad (13)$$

Comparing equation (8) with equation (13) reveals that the competition term (second terms in eqq. [8] and [13]) is larger under cultural transmission (weighted by $1 - m_{ct}$) than under genetic transmission (weighted by $(1 - m)^2$), provided the mixing rates (m_{ct} and m) are the same. This occurs because, unlike biological offspring, cultural

variants can spread to another group through imitation without reducing the number of copies of these variants in the source group. Hence, all else being equal, an increase in average group productivity, due to the expression of helping, increases local competition more under the cultural imitation model than under genetic transmission. This is also true when imitation occurs by pairwise payoff comparison, but since this situation is more complicated, we have included it in the appendix (see eqq. [A2]–[A8]). Finally, we mention that the assumption that $f_{ij} - f_j$ and $f_{ij} - f$ are of order δ does not imply that individuals have a very sensitive method of comparing payoffs. The difference between payoffs used to assess the highest payoff will depend on the difference between the cost C and the benefit B, which can be very different from each other, even when both are assumed to be small.

Results

Helping Affects Reproduction

We analyze the gradient of selection on the helping variant A when it affects the Darwinian fitness w. Substituting equation (8) into equation (4), using the payoff function (eq. [6]), and rearranging (see eqq. [A9]–[A15]), we find that the change in frequency of variant A is given to the first order of selection by

$$\Delta p = p(1 - p)[-C + BR - (1 - m)^2(B - C)R^R], \quad (14)$$

where R is the regression of the cultural variant at the helping locus of a randomly chosen group mate on the cultural variant in a focal individual (i.e., cultural relatedness; Allison 1991) and

$$R^R = \frac{1}{N} + \left(\frac{N - 1}{N}\right)R \quad (15)$$

is the cultural relatedness between a focal individual and an individual sampled with replacement from its group (i.e., including the focal individual with probability $1/N$).

The coefficient of cultural relatedness R predicts the tendency of a focal individual's group mate to express the cultural helping variant given that he/she carries that variant, and it corresponds here to the probability of identity in state between two homologous cultural variants sampled from two different individuals from the same group (see eqq. [A16], [A17]). This cultural relatedness will generally depend on selection, but because of the assumptions of weak selection and unbiased horizontal transmission, it is sufficient to evaluate it in a neutral model only ($\delta = 0$; Roze and Rousset 2003; Rousset 2004). This entails that R will be frequency-independent but will depend on

the life cycle parameters of the model, such as group size N, the migration rate m, and features of horizontal cultural transmission (see below). The term $p(1 - p)$ in equation (14) represents the variance in the frequency of A in the population, which does not affect the direction of selection, while the term in square brackets is the gradient of selection. Equation (14) informs us that the change in frequency p of variant A will depend on the balance of three terms: the direct cost C of the helping by a focal individual, the benefit BR received by a focal individual from all other helping individuals in his/her group, and the increase in competition faced by the focal individual's offspring due to the help expressed by him/her and other group members that augments group productivity. The selection pressure analogous to equation (14) was first obtained by Taylor (1992*a*) for genetically determined helping traits and can also be obtained as a special case of the selection pressure on culturally determined strong reciprocity in spatially subdivided populations (Lehmann et al. 2007, their eq. [A-17]).

The selection gradient on variant A depends on how actors affect their own fitness and the fitness of group neighbors (i.e., social effects). In order to identify the net effect $-c$ on its fitness of the behavior of a single focal individual expressing variant A, we set the cultural relatedness to 0 ($R = 0$) in the gradient of selection of equation (14), which removes all correlated social effects (Rousset 2004, chap. 7), giving

$$-c = -C - \frac{(1 - m)^2(B - C)}{N}. \quad (16)$$

Here, $-C$ is the direct fecundity cost to a focal individual expressing an act of helping, $(B - C)/N$ represents the increment in the average productivity of group members following the act of helping by the focal individual, and $(1 - m)^2$ is the probability that an offspring of the focal individual competes against another offspring produced in the focal group. Hence, by helping neighbors, the focal individual increments group productivity, which is deleterious for himself/herself because it decreases the likelihood that its own offspring will succeed. Following Hamilton (1964, 1970), an act of helping will subsequently be called "altruistic" whenever expressing variant A decreases the fitness of the actor while increasing the fitness of receptors when everything else is held constant. Thus, the cost of altruism is the difference in fitness between not expressing and expressing variant A while holding everything else constant, which gives c in equation (16) for all variant frequencies under our assumption of weak selection.

In the special case where the direction of selection on variant A is entirely accounted for by the net effect of an actor on its fitness, which occurs in the case of an isolated

panmictic population of size N (i.e., $m = 0$ and $R = 0$ in eqq. [14], [16]), it would be beneficial for the actor to harm neighbors ($B < 0$). This occurs because the resulting decrement of competition would increase the chance that the offspring of an actor who harms neighbors succeed, with the direction of selection on such a harming variant being positive (i.e., $-c > 0$) if $C/B < -1/(N - 1)$ (Hamilton 1971). But because the population is structured, the direction of selection on variant A is not determined solely by fitness effects on self but depends on the correlated behaviors of group members. With positive relatedness between group members ($R > 0$), more individuals are helping each other in groups, which, on the one hand, increases the intensity of competition and, on the other hand, leads to high indirect benefits. How do these opposite factors balance each other? From equation (14), the selection pressure on the helping variant A is positive at all variant frequencies ($\Delta p > 0$) if the cost to benefit ratio satisfies

$$\frac{C}{B} < \frac{[1 + m(2 - m)(N - 1)]R - (1 - m)^2}{N - (1 - m)^2[1 + (N - 1)R]}. \quad (17)$$

When the right-hand side of equation (17) is 0, the direction of selection on the helping variant is given by Taylor's $-C > 0$ rule (eq. [1]), and neither costly helping nor harming evolves. This occurs when the steady state cultural relatedness is given by the threshold value

$$R^* = \frac{(1 - m)^2}{1 + m(2 - m)(N - 1)}, \quad (18)$$

which is obtained by setting the right-hand side of equation (17) to 0 and solving for R. This value of relatedness corresponds to the relatedness obtained for genetic traits (i.e., Wright's measure of population structure; Wright 1931, 1951) under a Wright-Fisher mode of reproduction (i.e., all adults reproduce and die per unit time; Ewens 2004). At this value of R, the positive effect of relatedness on helping is exactly cancelled by the concomitant increase in competition; thus, the direction of selection on the trait is independent of the structure of the population (Taylor 1992a, 1992b). If the cultural relatedness is higher than the threshold value ($R > R^*$), the right-hand side of equation (17) is positive, and helping neighbors (i.e., $B > 0$) at a fecundity cost to self ($C > 0$) may be selected for. But if $R < R^*$, the cultural relatedness is lower than the threshold value, and the right-hand side of equation (17) will be negative. In that case, harming neighbors (i.e., $B < 0$) at a fecundity cost to self ($C > 0$) can be selected for.

For our model of unbiased horizontal cultural inheritance, the steady state value of the cultural relatedness is obtained from the recursion presented in the appendix (see eqq. [A19], [A20]) as

$$R = \frac{N\alpha_d + (1 - m)^2(\alpha_s - \alpha_d)}{N - (1 - m)^2(N - 1)(\alpha_s - \alpha_d)}, \quad (19)$$

where α_s denotes the probability that two different individuals from a group who carried identical cultural variants before horizontal transmission also have identical variants after horizontal transmission, and α_d denotes the probability that two individuals carrying different variants before horizontal transmission have identical cultural variants after cultural transmission. From this equation, it is easy to find a horizontal transmission scheme that will result in a cultural relatedness such that $R > R^*$ even in very large groups ($N \to \infty$), in which case $R = \alpha_d/[1 - (1 - m)^2(\alpha_s - \alpha_d)]$. Suppose that during horizontal transmission, each individual within a group adopts the cultural variant of a randomly chosen leader or teacher from its group with probability τ ("one to many" transmission; Cavalli-Sforza and Feldman 1981, chap. 3.11), while with complementary probability $1 - \tau$, an individual adopts the cultural variant from another individual sampled at random from its group. Then, we have $\alpha_d = \alpha_s = \tau^2$, and R is simply given by τ^2. Taking the limit $N \to \infty$ and $R = \tau^2$ in equation (17), we find

$$\frac{C}{B} < \frac{(2 - m)m\tau^2}{1 - (1 - m)^2\tau^2}, \quad (20)$$

which illustrates that if τ is sufficiently large, altruistic helping can invade the population under "one to many" transmission" in groups of any size even if migration is strong.

Helping Affects Cultural Preferences

We now analyze the gradient of selection on variant A when it affects the transmission function t. Substituting equation (13) into equation (5), using the payoff function (eq. [6]), assuming no genetic migration between groups ($m = 1$), and rearranging (see eqq. [A18], [A19]), we find that the change in frequency of variant A is given to the first order of selection by

$$\Delta p = p(1 - p)[-C + BR - (1 - m_{ct})(B - C)R^R], \quad (21)$$

which can also be obtained from Lehmann et al.'s (2007) equation [A-38] as a special case of the selection pressure on culturally determined strong reciprocity. The only difference between this equation and equation (14) is that the competition term is increased, as it is now weighted by the factor $1 - m_{ct}$ instead of $(1 - m)^2$. Consequently,

the net effect $-c$ of an individual bearing cultural variant A on its ability to transmit that variant to the next generation (obtained by setting $R = 0$ in eq. [21]) is given by

$$-c = -C - \frac{(1 - m_{ct})(B - C)}{N}. \tag{22}$$

Comparing this equation with equation (16), we see that the consequence of helping neighbors is more deleterious for an actor when cultural variants affect the nonreproductive transmission function t than when they affect the fitness function w. This suggests that the threshold value of cultural relatedness R^* above which helping instead of harming is selected for will be greater in the former case than in the latter. By following the same argument as for variants that are linked to physical reproduction, helping will subsequently be called "altruistic" whenever expressing variant A produces $-c < 0$.

From the selection gradient (eq. [21]), we find that the helping variant A is selected for at all variant frequencies if

$$\frac{C}{B} < \frac{[1 + m_{ct}(N - 1)]R - (1 - m_{ct})}{N - (1 - m_{ct})[1 + (N - 1)R]}. \tag{23}$$

Costly helping can now be selected for when the cultural relatedness exceeds the threshold value

$$R^* = \frac{1 - m_{ct}}{1 + m_{ct}(N - 1)}, \tag{24}$$

which is greater than that given by equation (18) for the same migration rates ($m_{ct} = m$). When the cultural relatedness takes this value, the direction of selection on the cultural variant is independent of population structure and is given by Taylor's $-C > 0$ rule. This corresponds to a Moran process of imitation, that is, when one individual in the population updates its strategy per unit time (Ewens 2004).

But what about the value of the cultural relatedness following our life cycle assumptions, where all individuals in the populations simultaneously update their strategy (i.e., Wright-Fisher process; Ewens 2004)? The cultural relatedness between two individuals under our imitation model is obtained as a solution to $R = (1 - m_{ct})^2\{1/N + [(N - 1)/N]R\}$ because with probability $(1 - m_{ct})^2$, two different individuals randomly sampled from a focal group both imitate the cultural variant of a focal group member, in which case they both imitate the same individual with probability $1/N$. Consequently, the cultural relatedness is given at steady state (see eq. [18]) by

$$R = \frac{(1 - m_{ct})^2}{1 + m_{ct}(2 - m_{ct})(N - 1)}. \tag{25}$$

Substituting equation (25) into equation (23), we find that the helping variant can invade the population when the cost to benefit ratio satisfies

$$\frac{C}{B} < \frac{-(1 - m_{ct})}{N(2 - m_{ct}) - (1 - m_{ct})}. \tag{26}$$

Since the right-hand side is always negative, only harming group mates (i.e., $B < 0$) can be selected for if there is a payoff cost ($C > 0$) to expressing variant A. Hence, if individuals tend to copy the cultural variant of group members under the present imitation scheme, selection tends to promote the evolution of traits decreasing the payoff to group members. Note that we do not label such a trait "spiteful" because harming is self-serving here (Foster et al. 2001). The cultural variant spreads only when its expression increases the number of copies of the actor's exemplar (i.e., $-c > 0$ with $-c$ given by eq. [22]).

Discussion

Cultural transmission occurs throughout the animal kingdom (Bonner 1980; Avital and Jablonka 2000) and is central to human sociality (Cavalli-Sforza and Feldman 1981; Lumsden and Wilson 1981; Richerson and Boyd 2005). Like genetic evolution, cultural evolution can be strongly affected by population subdivision, that is, cultural group selection, where differential group success results from the expression of different cultural traits (Fehr et al. 2002; Henrich 2004; Richerson and Boyd 2005). As with genetic evolution, selection on variants (predictors of phenotype) under cultural evolution (be they social or nonsocial) can be described in terms of direct and indirect selection coefficients weighted by statistical associations (e.g., linkage disequilibrium, relatedness coefficient) between variants, sampled at the same and/or different loci, from the same and/or from different individuals (e.g., Feldman and Cavalli-Sforza 1984; Feldman et al. 1985; Frank 1998; Roze and Rousset 2005; Gardner et al. 2007). But a key feature of cultural transmission relative to genetic transmission is the ease with which traits can be passed on to biologically unrelated individuals through oblique or horizontal transmission schemes, a process that can lead to high cultural relatedness within groups ("cultural inbreeding" [Cavalli-Sforza and Feldman 1981] or "cultural relatedness" [Allison 1991]). In this article, we have investigated how, through its effect on the dynamics of cultural relatedness, population subdivision influences the evolution of culturally inherited altruistic helping behaviors under different cultural transmission schemes. Our

analyses have revealed that cultural transmission can promote as well as inhibit the evolution of altruism relative to simple genetic transmission, thus leading to a variety of outcomes that depend on the way cultural transmission occurs.

When the expression of cultural variants directly affects the reproductive success of individuals, and when oblique and horizontal transmission are unbiased, selection on culturally determined helping traits operates qualitatively like selection on genetically determined helping (eq. [14]). The only difference occurs through effects on degree of cultural relatedness R in groups of interacting individuals. As in simple genetic models, this coefficient represents the probability that two individuals in a group display the same behavior; when it is high, altruistic helping can be favored because helping is directed toward other individuals bearing similar traits. We saw that cultural relatedness can be much greater than genetic relatedness (eq. [19]), especially when there is a "one to many" transmission rule in the population, where individuals tend to copy only a few individuals within groups who could be the "teachers" of the "leaders" of the group (Cavalli-Sforza and Feldman 1981). And this process is plausible: there is evidence that new innovations spread because a few "opinion leaders" take them up and are subsequently imitated by the rest of the group/village (Rogers 1995). It is also possible that "one to many" transmission schemes occur in nonhuman species, such as chimpanzees, where social learning is common and traditions—such as tool use, sexual behavior, and social organization—have been documented (Avital and Jablonka 2000). The prediction of cultural group selection (Fehr et al. 2002; Henrich 2004; Richerson and Boyd 2005), namely that cultural transmission can homogenize behaviors within groups and promote the evolution of altruistic helping in groups of very large size, is thus supported. But there is still the question of how nonvertical social learning rules can evolve in the first place. A recent model of the coevolution of helping and cultural transmission found that the benefits of helping are not sufficient to coselect for nonvertical transmission in groups of large size (Lehmann and Feldman, forthcoming), so that for social learning to evolve, other factors, which could be by-products of the evolution of individual learning in temporally varying environments, must be operating (Boyd and Richerson 1985; Wakano et al. 2004; Borenstein et al. 2008).

Our models that allow culturally determined helping to be transmitted by imitation/social learning with no effect of the strategy on reproduction offer a very different perspective from those in which helping affects reproduction. We considered a situation where individuals copy others from their group or from other groups and tend to copy individuals that perform well according to some arbitrary payoff or cultural criteria (so-called proportional updating). These payoffs could involve social standing or wealth for competing individuals and market share or the price of a good for competing firms. But in this case of nonreproductive differential success of cultural variants, cultural transmission may actually inhibit the evolution of helping behaviors. This occurs because, by incrementing the payoff to neighbors, an actor markedly decreases the likelihood of being imitated by other members of its group. The local increase in competition actually turns out to be stronger in the case where cultural variants affect nonreproductive differential success (cf. eqq. [7], [12]), which can be understood by recognizing the different role of "migration" under these two situations (fig. 1). When traits are linked to physical reproduction, migration of juveniles markedly reduces the intensity of local competition to which the offspring of a focal individual are exposed (see the denominator of the first term in eq. [7]; fig. 1c, reproductive column). By contrast, when "migration" occurs through an individual copying a member of another group, the transmission of a cultural trait to another group does not by itself reduce the amount of competition in the source group (see the denominator of the first term in eq. [12]), which results in a situation where a focal individual is exposed to increased competition stemming from his helping neighbors (fig. 1c, nonreproductive column). This difference turns out to be crucial under our life cycle assumptions because it may not only cancel out the payoff benefit of helping, as occurs in the classical models of Taylor (1992a, 1992b), but may go further and even select for the evolution of unconditional harming behavior toward group mates (eq. [26]). To our knowledge, this represents the first such example in a large subdivided population (formally infinite population size). Cultural transmission, therefore, may inhibit the evolution of altruistic helping.

Our imitation model also allowed us to identify life-history conditions under which the direction of selection on the cultural variant becomes independent of population structure, that is, when the direction of selection is given by the $-C > 0$ rule. This occurs under a Moran process of imitation with one individual in the population updating its strategy per unit time (see eq. [24]), whereas for genetic inheritance or when cultural variants affect reproduction, this condition is obtained when reproduction follows the Wright-Fisher process (see eq. [18]), where all individuals in the population simultaneously reproduce (see eq. [1]; Taylor 1992a, 1992b; Rousset 2004). It is well established that when helping affects Darwinian fitness and everything else remains equal, helping is more likely to be favored under "asynchronous" than "synchronous" updating (e.g., Taylor and Irwin 2000; Irwin and Taylor 2001; Hauert and Doebeli 2004). In line with these findings, our
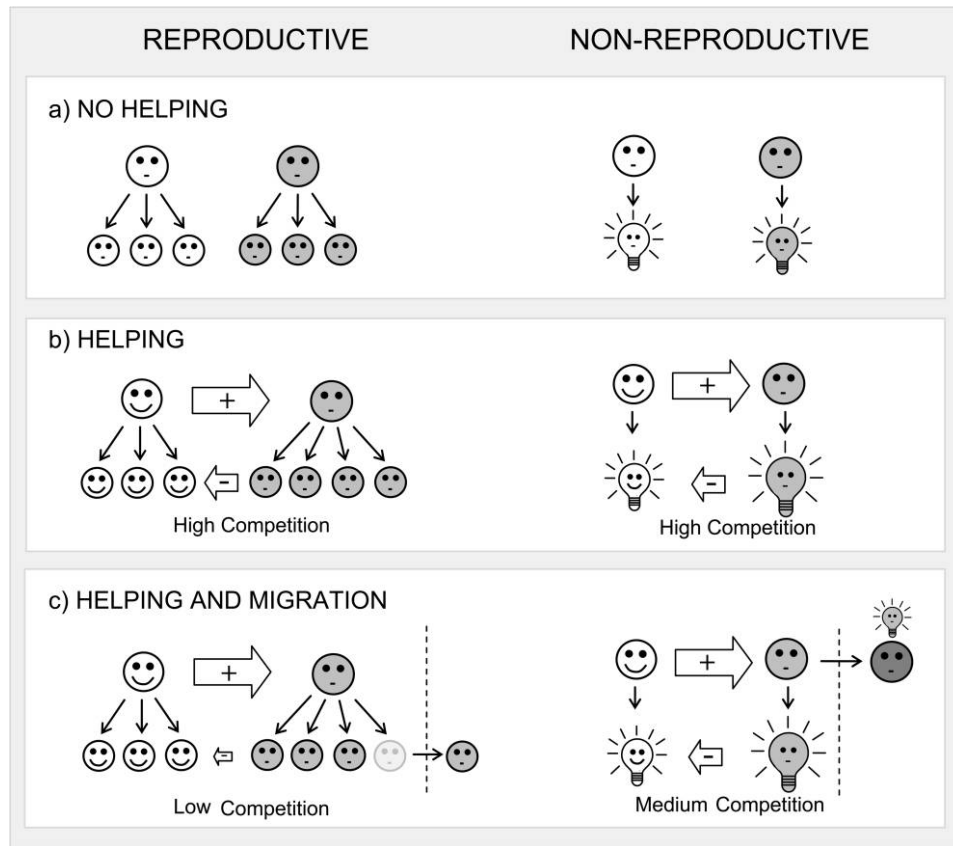
Figure 1: How cultural transmission can inhibit the evolution of helping behaviors, in cartoon form. Reproductive: the left column depicts a genetically inherited trait, or, equivalently, a cultural trait that propagates purely via reproductive effects. Nonreproductive: the right column depicts a cultural trait that propagates purely via nonreproductive effects; for example, helping means giving money, and individuals tend to copy the behavior of the wealthiest (the individual with the largest light bulb). *a*, No helping behaviors (neutral trait here). *b*, White helps gray (*large white arrow*) in groups where there is no migration. This increases gray's fitness or wealth but means that white experiences more local competition because of gray's success (*small white arrow*). For the reproductive case, this corresponds to altruistic helping. For the nonreproductive case, this is also altruism but with respect to the evolutionary currency (e.g., wealth) but not altruism in a reproductive sense. *c*, White helps gray in a group where there is "migration" of cultural variants between groups. When the variants are linked to physical reproduction, the migration rate m between groups means that only a proportion $(1 - m)$ of white's offspring remain in the focal group, which in turn face competition from only $1 - m$ of gray's offspring that remained philopatric (eq. [7]). These two terms combine, with the result that the net competitive influence of gray on white is approximatively of order $(1 - m)^2$ (eq. [8]). By contrast, when the variants are not linked to physical reproduction, the spread of ideas between groups at rate m_{ct} does not involve offspring physically leaving the group. Hence, conditional on white copying an individual locally (with probability $1 - m_{ct}$), the competitive effect of gray on white resulting from white helping gray is not reduced (eq. [12]). This implies that the net competitive influence of gray on white is of order $1 - m_{ct}$, not $(1 - m_{ct})^2$ (eq. [13]). This may result in the evolution of self-serving harming behaviors rather than altruistic helping.

results suggest that under an imitation model, harming is more likely to be favored under "synchronous" than "asynchronous" updating. Will cultural processes that act independently of reproduction always impede the spread of helping? Probably not; if we assume that imitation always occurs on the basis of payoffs relative to the average payoff of the population, then helping would have no effect on local competition and would evolve more easily. This scenario is in fact analogous to the process of "selective emigration" proposed by Rogers (1990) for genetically in-

herited helping, a process implying that competition occurs only among emigrants globally and that groups with more helpers produce more emigrants. Since there is no effect of helping on local competition under this scenario, helping can invade the population when the cost to benefit ratio satisfies $C/B < R$, whether variant *A* affects Darwinian fitness or nonreproductive payoffs. Nevertheless, there is empirical evidence that individuals may prefer local information over global information when copying others (Rogers 1995). While it remains unclear whether

cultural evolution occurs without reproductive effects in nonhuman species, evidence of social comparison abounds in the human lineage (Richerson and Boyd 2005; Fliessbach et al. 2007) and is frequently used in assumptions used to construct models of economic agents (e.g., Ellison 1993; Kandori et al. 1993; Young 1993).

Previous models of cultural transmission of social traits resulting in nonreproductive differential success have sometimes assumed that imitation occurs by "pairwise payoff comparison" rules (Boyd and Richerson 2002; Boyd et al. 2003). In our results, such pairwise imitation schemes result in qualitatively similar dynamics to our "group payoff comparison" rule (cf. eqq. [13], [A8]), namely harming rather than helping behaviors are favored. This result seems to be at variance with the view that group-beneficial norms can spread rapidly as a result of cultural group selection when transmission follows these imitation schemes (Boyd and Richerson 2002; Boyd et al. 2003; Henrich 2004). There are two main reasons for the difference between these results. The first is that Boyd and Richerson (2002) assume that groups are of infinite size ($N \to \infty$), which removes any self-serving benefits from harming neighbors (i.e., the second term in eq. [22] vanishes). In that situation, selection favors helping only when $-C > 0$. Another difference, however, is that the social norm modeled by Boyd and Richerson (2002) is costly only as long as the frequency of the norm in groups is below a threshold frequency p^*; otherwise, expressing the social norm is beneficial for the actor. They assume that there is a selfish benefit of adopting the norm once many individuals express it. Such a social norm cannot invade when rare in groups of infinite size, and Boyd and Richerson (2002) assume that this social threshold frequency is reached by initially fixing the social norm in at least one group in the population, although this requires additional mechanisms, since cultural drift cannot produce fixation of the norm in groups of infinite size.

In the appendix, we derive Boyd and Richerson's model for finite group size by following exactly their assumptions concerning the costs and benefits resulting from the expression of the social norm (see eqq. [A21]–[A34]). In this case, the social norm to help others can be favored even when rare because, as a result of the buildup of cultural relatedness in groups of finite size by cultural drift, an actor is likely to interact with other individuals bearing the same social norm. Thus, finite group size allows the social threshold frequency p^* to be crossed naturally. The condition under which selection favors the norm will ultimately depend on group size and the cultural mixing rate; in our model, this condition is met for large group size and in the absence of a public good benefit when $p^* < 1/[Nm_{ct}(2 - m_{ct}) - (1 - m_{ct})]$ (see eq. [A34]). But had we assumed genetic inheritance, the selective pressure on the norm would have increased because, everything

else being equal, the increase in competition resulting from the expression of a social trait is higher in the model of Boyd and Richerson (2002) than under genetic transmission (cf. eqq. [8], [13]). This illustrates that our model is fully consistent with the results found by Boyd and Richerson (2002), but the different conclusions stem from modeling different social traits. We considered that helping is altruistic at all variant frequency, while Boyd and Richerson (2002) considered that helping is altruistic only below a threshold frequency, otherwise the social norm is self-serving.

We close then with what in hindsight may seem to be an obvious point. Just because we humans have cultural transmission, it should not be assumed that this leads to a more cooperative or altruistic state than exists in those species that lack it. Indeed, our models suggest that it will be those cultural transmission schemes that become the most detached from biological reproduction ("pure cultural transmission"; Lumsden and Wilson 1980; Lumsden and Wilson 1981) that are the most likely to result in selection for harming behavior.

Acknowledgments

Literature Cited

Allison, P. D. 1991. Cultural relatedness under oblique and horizontal transmission. Ethology and Sociobiology 13:153–169.

Aoki, K. 1982. A condition for group selection to prevail over counteracting individual selection. Evolution 36:832–842.

Avital, E., and E. Jablonka. 2000. Animal traditions. Cambridge University Press, Cambridge.

Axelrod, R., R. A. Hammond, and A. Grafen. 2004. Altruism via kin-selection strategies that rely on arbitrary tags with which they coevolve. Evolution 58:1833–1838.

Bonner, J. T. 1980. The evolution of culture in animals. Princeton University Press, Princeton, NJ.

Borenstein, E., K. Aoki, and M. W. Feldman. 2008. Evolution of learning in fluctuating environments: when selection favors both social and exploratory individuals. Evolution 62:573–602.

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

Boyd, R., and P. J. Richerson. 1985. Culture and the evolutionary process. University of Chicago Press, Chicago.

———. 2002. Group beneficial norms can spread rapidly in a structured population. Journal of Theoretical Biology 215:287–296.

———. 2005. The origin and evolution of cultures. Oxford University Press, Oxford.

Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson. 2003. The evolution of altruistic punishment. Proceedings of the National Academy of Sciences of the USA 100:3531–3535.

Bürger, R. 2000. The mathematical theory of selection, recombination, and mutation. Wiley, New York.

Cavalli-Sforza, L., and M. W. Feldman. 1973. Models for cultural inheritance. I. Group mean and within group variation. Theoretical Population Biology 4:42–55.

———. 1981. Cultural transmission and evolution. Princeton University Press, Princeton, NJ.

Cherry, J. L., and J. Wakeley. 2003. A diffusion approximation for selection and drift in a subdivided population. Genetics 163:421–428.

Crow, J. F., and K. Aoki. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. Proceedings of the National Academy of Sciences of the USA 81:6073–6077.

Ellison, G. 1993. Learning, local interaction, and coordination. Econometrica 61:1047–1071.

Eshel, I. 1972. On the neighbor effect and the evolution of altruistic traits. Theoretical Population Biology 11:258–277.

Ewens, W. J. 2004. Mathematical population genetics. Springer, New York.

Fehr, E., U. Fischbacher, and S. Gachter. 2002. Strong reciprocity, human cooperation, and the enforcement of social norms. Human Nature 13:1–25.

Feldman, M. W., and L. Cavalli-Sforza. 1984. Cultural and biological evolutionary processes: gene-culture disequilibrium. Proceedings of the National Academy of Sciences of the USA 81:1604–1607.

Feldman, M. W., L. Cavalli-Sforza, and J. L. Peck. 1985. Gene-culture coevolution: models for the evolution of altruism with cultural transmission. Proceedings of the National Academy of Sciences of the USA 82:5814–5818.

Fliessbach, K., B. Weber, P. Trautner, T. Dohem, U. Sunde, C. E. Elger, and A. Falk. 2007. Social comparison affects reward-related brain activity in the human ventral striatum. Science 318:1305–1308.

Foster, K. R., T. Wenseleers, and F. L. W. Ratnieks. 2001. Spite: Hamilton's unproven theory. Annales Zoologici Fennici 38:229–238.

Frank, S. A. 1998. Foundations of social evolution. Princeton University Press, Princeton, NJ.

Gardner, A., and S. A. West. 2006. Demography, altruism, and the benefits of budding. Journal of Evolutionary Biology 19:1707–1716.

Gardner, A., S. A. West, and N. Barton. 2007. The relation between multilocus population genetics and social evolution theory. American Naturalist 169:207–226.

Goodall, J. 1966. Tool using and aimed throwing in a community of free-living chimpanzees. Nature 201:1264–1266.

Grafen, A. 2007. An inclusive fitness analysis of a network model of Ohtsuki et al. Journal of Evolutionary Biology 20:2278–2283.

Hamilton, W. D. 1964. The genetical evolution of social behaviour. I. Journal of Theoretical Biology 7:1–16.

———. 1970. Selfish and spiteful behavior in an evolutionary model. Nature 228:1218–1220.

———. 1971. Selection of selfish and altruistic behaviour in some extreme models. Pages 59–91 *in* J. Eisenberg and W. Dillon, eds. Man and beast: comparative social behavior. Smithsonian Institution, Washington, DC.

Hauert, C., and M. Doebeli. 2004. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. Nature 428:643–646.

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

Henrich, J., and R. Boyd. 2002. On modeling cognition and culture: why cultural evolution does not require replication of representations. Journal of Cognition and Culture 2:87–112.

Hofbauer, J., and K. Sigmund. 1998. Evolutionary games and population dynamics. Cambridge University Press, Cambridge.

Irwin, A. J., and P. D. Taylor. 2001. Evolution of altruism in stepping-stone populations with overlapping generations. Theoretical Population Biology 60:315–325.

Jansen, V. A. A., and M. van Baalen. 2006. Altruism through beard chromodynamics. Nature 440:663–666.

Kandori, M., G. Mailath, and R. Rob. 1993. Learning, mutation, and long run equilibria in games. Econometrica 61:29–56.

Kendal, J., M. W. Feldman, and K. Aoki. 2006. Cultural coevolution of norm adoption and enforcement when punishers are rewarded or non-punishers are punished. Theoretical Population Biology 70:10–25.

Kirk, K. M., S. P. Blomberg, D. L. Duffy, A. C. Heath, I. P. Owens, and N. G. Martin. 2001. Natural selection and quantitative genetics of life-history traits in Western women: a twin study. Evolution 55:423–435.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727–1750.

Le Galliard, J., R. Ferrière, and U. Dieckmann. 2003. The adaptive dynamics of altruism in spatially heterogeneous populations. Evolution 57:1–17.

Lehmann, L. 2006. The evolution of trans-generational altruism: kin selection meets niche construction. Journal of Evolutionary Biology 20:181–189.

Lehmann, L., and F. Balloux. 2007. Natural selection on fecundity variance in subdivided populations: kin selection meets bet hedging. Genetics 176:361–377.

Lehmann, L., and M. W. Feldman. Forthcoming. The coevolution of altruism and cultural transmission under random group formation. Theoretical Population Biology.

Lehmann, L., N. Perrin, and F. Rousset. 2006. Population demography and the evolution of helping behaviors. Evolution 60:1137–1151.

Lehmann, L., F. Rousset, D. Roze, and L. Keller. 2007. Strong reciprocity or strong ferocity? a population genetic view of the evolution of altruistic punishment. American Naturalist 170:21–36.

Lumsden, C. J., and E. O. Wilson. 1980. Translation of epigenetic rules of individual behavior into ethnographic patterns. Proceedings of the National Academy of Sciences of the USA 77:4382–4386.

———. 1981. Genes, mind and culture. Harvard University Press, Cambridge, MA.

Nagylaki, T. 1993. The evolution of multilocus systems under weak selection. Genetics 134:627–647.

Ohtsuki, H., C. Hauert, E. Lieberman, and M. A. Nowak. 2006. A simple rule for the evolution of cooperation on graphs and social networks. Nature 441:502–505.

Price, G. R. 1970. Selection and covariance. Nature 227:520–521.

Richerson, P. J., and R. Boyd. 2005. Not by genes alone. University of Chicago Press, Chicago.

Rogers, A. R. 1990. Group selection by selective emigration: the effects of migration and kin structure. American Naturalist 135:398–413.

Rogers, M. E. 1995. Diffusion of innovations. 4th ed. Free Press, New York.

Rousset, F. 2002. Inbreeding and relatedness coefficients: what do they measure? Heredity 88:371–380.

———. 2004. Genetic structure and selection in subdivided populations. Princeton University Press, Princeton, NJ.

———. 2006. Separation of time scales, fixation probabilities and convergence to evolutionarily stable states under isolation by distance. Theoretical Population Biology 69:165–179.

Rousset, F., and D. Roze. 2007. Constraints on the origin and maintenance of genetic kin recognition. Evolution 61:2320–2330.

Roze, D., and F. Rousset. 2003. Selection and drift in subdivided populations: a straightforward method for deriving diffusion approximations and applications involving dominance, selfing and local extinctions. Genetics 165:2153–2166.

———. 2005. Inbreeding depression and the evolution of dispersal rates: a multilocus model. American Naturalist 166:708–721.

Taylor, P. D. 1992*a*. Altruism in viscous populations: an inclusive fitness model. Evolutionary Ecology 6:352–356.

———. 1992*b*. Inclusive fitness in a homogeneous environment. Proceedings of the Royal Society B: Biological Sciences 240:299–302.

Taylor, P. D., and A. J. Irwin. 2000. Overlapping generations can promote altruistic behavior. Evolution 54:1135–1141.

van Baalen, M., and A. Rand. 1998. The unit of selection in viscous populations and the evolution of altruism. Journal of Theoretical Biology 193:631–648.

van Veelen, M., and A. Hopfensitz. 2007. In love and war: altruism, norm formation, and two different types of group selection. Journal of Theoretical Biology 249:667–680.

Wakano, J. Y. 2007. Evolution of cooperation in spatial public goods games with common resource dynamics. Journal of Theoretical Biology 247:616–622.

Wakano, J. Y., K. Aoki, and M. W. Feldman. 2004. Evolution of social learning: a mathematical analysis. Theoretical Population Biology 66:249–258.

West, S. A., A. S. Griffin, and A. Gardner. 2007. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. Journal of Evolutionary Biology 20:415–432.

Whitlock, M. C. 2002. Selection, load and inbreeding depression in a large metapopulation. Genetics 160:1191–1202.

Wilson, D., G. Pollock, and L. Dugatkin. 1992. Can altruism evolve in purely viscous populations? Evolutionary Ecology 6:331–341.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97–159.

———. 1951. The genetical structure of populations. Annals of Eugenics 15:323–354.

Young, H. P. 1993. The evolution of conventions. Econometrica 61:57–84.

Appendix from L. Lehmann et al., "Cultural Transmission Can Inhibit the Evolution of Altruistic Helping"

(Am. Nat., vol. 172, no. 1, p. 12)

Evaluation of the Selection Gradients

Pairwise Payoff Comparison Rule

We derive here the transmission function t_{ij} for the model where imitation occurs with a "pairwise payoff comparison" rule. Under this transmission rule, with probability $1 - m_{ct}$, a focal individual compares his/her payoff with that of a randomly chosen individual from his/her group (excluding himself/herself) and adopts the variant of the group mate with a probability given by the payoff to that individual relative to the sum of the focal individual's and the group mate's payoffs. With complementary probability m_{ct}, the focal individual compares his/her payoff with that of an individual chosen at random from another group (nonfocal group) and adopts the variant of that individual with a probability given by the payoff to that individual relative to the sum of the focal individual's and the chosen individual's payoffs. This updating rule is equivalent to the one considered by Boyd and Richerson (2002), Boyd et al. (2003), and Lehmann et al. (2007), although Boyd and Richerson (2002) considered a more general population structure where there is isolation by distance, but they assumed infinite group size. Here we consider the effect of finite group size but without isolation by distance. From the assumptions just described, the probability that individual i from group j retains its cultural variant is

$$\Pr(ij \leftarrow ij) = \sum_{k \neq i} \frac{(1 - m_{ct})(1 + f_{ij})}{(N - 1)[(1 + f_{ij}) + (1 + f_{kj})]} + \sum_{l \neq j} \sum_{k} \frac{m_{ct}(1 + f_{ij})}{(n_d - 1)N[(1 + f_{ij}) + (1 + f_{kl})]}, \tag{A1}$$

while the probability that individual k ($k \neq i$) from group j adopts the cultural variant of individual i from group j is

$$\Pr(kj \leftarrow ij) = \frac{(1 - m_{ct})(1 + f_{ij})}{(N - 1)[(1 + f_{ij}) + (1 + f_{kj})]}. \tag{A2}$$

Finally, the probability that individual k from group l adopts the cultural variant of individual i from group j is

$$\Pr(kl \leftarrow ij) = \frac{m_{ct}(1 + f_{ij})}{(n_d - 1)N[(1 + f_{ij}) + (1 + f_{kl})]}. \tag{A3}$$

By inserting these equations into the transmission function (eq. [9]) expanded as

$$t_{ij} = \Pr(ij \leftarrow ij) + \sum_{k \neq i} \Pr(kj \leftarrow ij) + \sum_{l \neq j} \sum_{k} \Pr(kl \leftarrow ij), \tag{A4}$$

we obtain

$$t_{ij} = (1 - m_{ct})\left[\frac{1}{N - 1} \sum_{k \neq i} \frac{2(1 + f_{ij})}{(1 + f_{ij}) + (1 + f_{kj})} \right] + m_{ct}\left[\frac{1}{(n_d - 1)N} \sum_{l \neq j} \sum_{k} \frac{2(1 + f_{ij})}{(1 + f_{ij}) + (1 + f_{kl})} \right]. \tag{A5}$$

Under weak selection, $2(1 + f_{ij})/[(1 + f_{ij}) + (1 + f_{kl})] \simeq 1 + (f_{ij} - f_{kl})/2$, and using $[1/(n_d - 1)N] \sum_{l \neq j} \sum_{k} f_{kl} \to f$ as $n_d \to \infty$ allows us to write

1

$$t_{ij} \simeq 1 + \frac{(1 - m_{\text{ct}})}{2}\left(f_{ij} - \frac{1}{N-1}\sum_{k\neq i}f_{kj}\right) + \frac{m_{\text{ct}}}{2}(f_{ij} - f). \tag{A6}$$

Noting that

$$\frac{1}{N-1}\sum_{k\neq i}f_{kj} = \frac{Nf_j}{N-1} - \frac{f_{ij}}{N-1}, \tag{A7}$$

we find after rearrangement that the transmission function is given by

$$t_{ij} \simeq 1 + \frac{1}{2}\left[(f_{ij} - f) - (1 - m_{\text{ct}})\left(\frac{Nf_j}{N-1} - \frac{f_{ij}}{N-1} - f\right)\right]. \tag{A8}$$

Comparing this equation with equation (13), we see that the increase in local competition is slightly lower under "pairwise payoff" comparison than under "group payoff" comparison but remains much larger than under genetic transmission (cf. eqq. [8], [A8]). The factor 1/2 in front of the square bracket will not affect the direction of selection, only the speed of evolution, and when group size becomes large ($N \rightarrow \infty$), the term in square brackets in equation (A8) converges to the second term of equation (13).

Selection Gradient

Darwinian Model

We derive here the gradient of selection on variant A (eq. [14]) when it affects the fitness function w. To that end, we express equation (6) as

$$1 + f_{ij} = 1 - Cp_{ij} + Bp_{-ij}, \tag{A9}$$

where $p_{-ij} = [1/(N-1)]\sum_{k\neq i}p_{kj}$ designates the average frequency of variant A in group j but excluding individual i from the average. Averaging equation (A9) over all individuals within group j, we have

$$1 + f_j = \frac{1}{N}\sum_i(1 + f_{ij}),$$
$$= 1 + (B - C)p_j, \tag{A10}$$

where $p_j = 1/N\sum_i p_{ij}$ is the average frequency of A in group j. Finally, the average of equation (A10) over all groups in the population is

$$1 + f = \frac{1}{n_{\text{d}}}\sum_j(1 + f_j),$$
$$= 1 + (B - C)p, \tag{A11}$$

where p is the average frequency of A in the population and n_{d} is the number of demes ($n_{\text{d}} \rightarrow \infty$ due to our assumption of the infinite island model of dispersal). Substituting equations (A9)–(A11) into equation (8) yields to the first-order effect of selection:

$$w_{ij} = 1 - Cp_{ij} + Bp_{-ij} - (1 - m)^2(B - C)p_j - [1 - (1 - m)^2](B - C)p. \tag{A12}$$

Inserting this fitness function into equation (4) and noting that the covariance of a sum can be decomposed into a sum of covariances and that the covariance between a constant and a random variable is 0, we obtain

$$\Delta p = -C\,\mathrm{Cov}\,(p_{ij}, p_{ij}) + B\,\mathrm{Cov}\,(p_{-ij}, p_{ij}) - (1-m)^2(B-C)\,\mathrm{Cov}\,(p_j, p_{ij}) - [1-(1-m)^2](B-C)\,\mathrm{Cov}\,(p, p_{ij}).$$

$$(A13)$$

Because of our assumption of the infinite island model of dispersal, the covariance $\mathrm{Cov}\,(p, p_{ij})$ between the variant carried by individual i from group j and a homologous variant sampled in one individual chosen at random from the population is 0. Since $\mathrm{Cov}\,(p_{ij}, p_{ij}) = p(1-p)$ is the genetic variance in the population and the average frequency of variant A in group j can be written as

$$p_j = \frac{1}{N}p_{ij} + \left(\frac{N-1}{N}\right)\frac{1}{N-1}\sum_{k\neq i}p_{kj},$$

$$= \frac{1}{N}p_{ij} + \left(\frac{N-1}{N}\right)p_{-ij}, \qquad (A14)$$

we can factor out $p(1-p)$ from equation (A13) and rearrange it to obtain

$$\Delta p = p(1-p)[-C + BR - (1-m)^2(B-C)R^{\mathrm{R}}], \qquad (A15)$$

where the term in square brackets is the gradient of selection on variant A,

$$R = \frac{\mathrm{Cov}\,(p_{-ij}, p_{ij})}{p(1-p)} \qquad (A16)$$

is the regression of the cultural variant at the helping locus of group mates on the variant at the helping locus of a focal individual (i.e., cultural relatedness), and $R^{\mathrm{R}} = [1 + (N-1)R]/N$.

From the definition of the covariance between two random variables (i.e., $\mathrm{Cov}\,(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$) and from equation (A16), the probability $\mathrm{E}[p_{ij}p_{-ij}]$ that two homologous cultural variants sampled without replacement are both A can be written as

$$\mathrm{E}[p_{ij}p_{-ij}] = Rp + (1-R)p^2. \qquad (A17)$$

This equation suggests that R can also be interpreted as the probability that the lineages of two homologous variants sampled in two different individuals descend from the same recent common cultural ancestor (cumulative probability of coalescence in genetics; see Rousset 2002; Roze and Rousset 2003), in which case both individuals carry allele A with probability p. With complementary probability $1-R$, the lineages of the two variants have not coalesced in the recent past (they do not descend from the same recent common ancestor), in which case the two lineages are considered as independent and the two individuals bear variant A with probability p^2. This interpretation of the kinship coefficient R in terms of probabilities of identity between variants carried by interacting individuals is valid whenever the kinship associations reach their steady state before any significant change in allele frequency has occurred at the level of the population, which implies that selection must be weaker than the other process affecting the dynamics of kinship. This "quasi equilibrium" assumption is often used in population genetic theory (e.g., Nagylaki 1993; Bürger 2000; Kirkpatrick et al. 2002; Whitlock 2002; Cherry and Wakeley 2003; Roze and Rousset 2003; Rousset 2006) and allows us to conveniently evaluate the relatedness in terms of model parameters (see "Cultural Relatedness for the Darwinian Fitness Model").

Group Payoff Comparison

We now carry out an analysis similar to that used in "Darwinian Model" in order to derive the selection gradient on variant A, when it affects the transmission function t (eq. [21]). Substituting equations (A9)–(A11) into equation (13) gives to the first order effect of selection

$$t_{ij} = 1 - Cp_{ij} + Bp_{-ij} - (1-m_{\mathrm{ct}})(B-C)p_j - m_{\mathrm{ct}}(B-C)p. \qquad (A18)$$

Inserting this equation into equation (5) and using the results of "Darwinian Model," we obtain under weak section

$$\Delta p = -C \operatorname{Cov}(p_{ij}, p_{ij}) + B \operatorname{Cov}(p_{-ij}, p_{ij}) - (1 - m_{ct})(B - C) \operatorname{Cov}(p_j, p_{ij}),$$

$$= p(1 - p)[-C + BR - (1 - m_{ct})(B - C)R^R]. \tag{A19}$$

Cultural Relatedness for the Darwinian Fitness Model

Here, we present the equilibrium value of the cultural relatedness coefficient R under unbiased horizontal cultural transmission for the Darwinian fitness model. To that end, we define two parameters: α_s denotes the probability that two individuals sampled from a group and bearing identical cultural variants before the stage of cultural transmission (stage 4 of the life cycle) bear an identical cultural variant after cultural transmission, and α_d denotes the probability that two individuals bearing different cultural variants before the stage of cultural transmission (stage 4 of the life cycle) bear the identical cultural variant after cultural transmission. With these two parameters, the probability that two individuals sampled in a group after cultural transmission bear the same cultural variant identical in state is given at steady state by

$$R = (1 - m)^2 \left\{ \frac{1}{N}\alpha_s + \left(\frac{N - 1}{N}\right)[R\alpha_s + (1 - R)\alpha_d] \right\} + [1 - (1 - m)^2]\alpha_d, \tag{A20}$$

where $(1 - m)^2$ is the probability that two individuals stem from the same group and $1/N$ is the probability that they descend from the same parent. This equation is formally similar to the classical recursions for identity in state under a model of recurrent mutation (Crow and Aoki 1984), and its solution is

$$R = \frac{N\alpha_d + (1 - m)^2(\alpha_s - \alpha_d)}{N - (1 - m)^2(N - 1)(\alpha_s - \alpha_d)}. \tag{A21}$$

Spread of Group Beneficial Norms

We derive here the gradient of selection on a group beneficial norm (variant A) by assuming exactly the same payoff structure as the one used by Boyd and Richerson (2002, p. 290). Using our notation, the payoff to individual i in group j (relative to the payoff of an individual that did not give or receive benefits from the expression of the social norm) is written as

$$1 + f_{ij} = 1 + sp_{ij}(p_j - p^*) + \tilde{B}p_j, \tag{A22}$$

where $\tilde{B}$ is a public good benefit (shared equally by everybody in the group) resulting from the expression of the social norm and s is a benefit to individual i in group j expressing the social norm, whenever the average frequency p_j of the norm in group j exceeds the threshold frequency p^*; otherwise, it is costly to express the social norm. Averaging the payoff function over all individuals within group j, we have

$$1 + f_j = \frac{1}{N}\sum_i (1 + f_{ij}),$$

$$= 1 + sp_j(p_j - p^*) + \tilde{B}p_j, \tag{A23}$$

and the average payoff in the whole population is

$$1 + f = \frac{1}{n_{\mathrm{d}}N}\sum_j \sum_i (1 + f_{ij}),$$

$$= \frac{1}{n_{\mathrm{d}}N}\sum_j \sum_i (1 + s p_{ij} p_j - s p_{ij} p^* + \tilde{B} p_j), \tag{A24}$$

$$= 1 + s\mathrm{E}[p_j^2] - s p p^* + \tilde{B}p.$$

With these payoffs functions, we can evaluate the gradient of selection on the social norm under both genetic inheritance and cultural inheritance and, when cultural transmission follows the "Darwinian" model, the "pairwise payoff" and "group payoff" comparison models. We will consider here only the "group payoff" comparison model because it is qualitative similar but simpler that the "pairwise payoff" comparison model, which was originally considered by Boyd and Richerson (2002).

Under the "group payoff" comparison model, the transmission function (eq. [13]) for the social norm is obtained by using the payoff functions (eqq. [A21]–[A23]), which gives to first order

$$t_{ij} = 1 + s(p_{ij} p_j - p_{ij} p^*) + \tilde{B} p_j - (1 - m_{\mathrm{ct}})[s(p_j^2 - p_j p^*) + \tilde{B} p_j] - m_{\mathrm{ct}}\{s(\mathrm{E}[p_j^2] - p p^*) + \tilde{B}p\}. \tag{A25}$$

Substituting equation (A25) into equation (5) and using $\mathrm{Cov}\,(p, p_{ij}) = 0$ and $\mathrm{Cov}\,(\mathrm{E}[p_j^2], p_{ij}) = 0$, we obtain

$$\Delta p = s[\mathrm{Cov}\,(p_{ij} p_j, p_{ij}) - p^*\,\mathrm{Cov}\,(p_{ij}, p_{ij})] + \tilde{B}\,\mathrm{Cov}\,(p_j, p_{ij})$$

$$- (1 - m_{\mathrm{ct}})\{s[\mathrm{Cov}\,(p_j^2, p_{ij}) - p^*\,\mathrm{Cov}\,(p_j, p_{ij})] + \tilde{B}\,\mathrm{Cov}\,(p_j, p_{ij})\}. \tag{A26}$$

Since p_{ij} is an indicator variable equal to 1 or 0, we have $p_{ij} = p_{ij}^2$, so that $\mathrm{E}[p_j p_{ij} p_{ij}] = \mathrm{E}[p_j p_{ij}]$. Then, from $\mathrm{Cov}\,(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$, we have

$$\mathrm{Cov}\,(p_j p_{ij}, p_{ij}) = \mathrm{E}[p_j p_{ij} p_{ij}] - \mathrm{E}[p_j p_{ij}]\mathrm{E}[p_{ij}],$$

$$= \mathrm{E}[p_j p_{ij}](1 - p), \tag{A27}$$

because $\mathrm{E}[p_{ij}] = p$. The probability $\mathrm{E}[p_j p_{ij}]$ that two homologous cultural variants randomly sampled with replacement from the same group are both A is equal to $\mathrm{E}[p_j^2]$ ($\mathrm{E}[p_j^2] = \sum_j p_j^2/n_{\mathrm{d}}$ and $\mathrm{E}[p_j p_{ij}] = \sum_{i,j} p_j p_{ij}/(n_{\mathrm{d}}N)$) and can be written in the infinite island model as

$$\mathrm{E}[p_j^2] = R^{\mathrm{R}} p + (1 - R^{\mathrm{R}})p^2. \tag{A28}$$

Using this equation, we have

$$\mathrm{Cov}\,(p_j p_{ij}, p_{ij}) = [R^{\mathrm{R}} + (1 - R^{\mathrm{R}})p]p(1 - p). \tag{A29}$$

The definition of the covariance also gives

$$\mathrm{Cov}\,(p_j^2, p_{ij}) = \mathrm{E}[p_j^2 p_{ij}] - \mathrm{E}[p_j^2]\mathrm{E}[p_{ij}], \tag{A30}$$

where $\mathrm{E}[p_j^2 p_{ij}]$ is the probability that three homologous cultural variants randomly sampled with replacement from the same group are all of type A ($\mathrm{E}[p_j^2 p_{ij}] = \mathrm{E}[p_j^3]$). This is

$$\mathrm{E}[p_j^3] = R_3^{\mathrm{R}} p + R^{\mathrm{R}} p^2 + (1 - R^{\mathrm{R}} - R_3^{\mathrm{R}})p^3, \tag{A31}$$

where R_3^{R} can be interpreted as the probability that the lineages of three homologous variants sampled in three different individuals all descend from the same recent cultural ancestor (see Roze and Rousset 2003). Combining the last two expressions, we obtain

$$\frac{\mathrm{Cov}\,(p_j^2, p_{ij})}{p(1-p)} = (1+p)R_3^{\mathrm{R}}. \tag{A32}$$

Taking into account all terms, the change in frequency of the social norm under weak selection is

$$\Delta p = p(1-p)\{s[R^{\mathrm{R}} + (1 - R^{\mathrm{R}})p - p^*] + m_{\mathrm{ct}}\tilde{B}R^{\mathrm{R}} - (1 - m_{\mathrm{ct}})s[(1+p)R_3^{\mathrm{R}} - p^*R^{\mathrm{R}}]\}, \tag{A33}$$

where the gradient of selection is now frequency dependent. In order to obtain an expression for Δp in closed form, it now remains to evaluate the probabilities of identity R^{R} and R_3^{R}. The probability R^{R} is obtained by substituting the solution of equation (25) into equation (15), while the recursion for R_3^{R} can be written down (e.g., Roze and Rousset 2003), but its solution is a bit complicated. In order to get around this problem, we now assume large group size. In that case, we can neglect R_3^{R} because $R_3^{\mathrm{R}} = O(1/N^2)$, where $O(1/N^2)$ is a remainder of order $1/N^2$, which represents the probability that two individuals have copied a cultural variant from the same individual and gives the rate at which R_3^{R} builds up per generation. To the first order in group size (neglecting all terms of order $1/N^2$ and of higher order), the change in variant frequency is given by

$$\Delta p = p(1-p)\{s[R^{\mathrm{R}} + (1 - R^{\mathrm{R}})p - p^*] + [m_{\mathrm{ct}}\tilde{B} + (1 - m_{\mathrm{ct}})sp^*]R^{\mathrm{R}}\}. \tag{A34}$$

With this assumption of large group size, the social norm can invade the population when rare ($p \to 0$) if the threshold frequency p^* satisfies the inequality

$$p^* < \frac{R^{\mathrm{R}}(s + m_{\mathrm{ct}}\tilde{B})}{[1 - (1 - m_{\mathrm{ct}})R^{\mathrm{R}}]s}, \tag{A35}$$

where $R^{\mathrm{R}} = 1/[Nm_{\mathrm{ct}}(2 - m_{\mathrm{ct}})]$, which is obtained by substituting equation (25) into equation (15) and assuming large group size. If $\tilde{B} = 0$, the inequality simplifies to $p^* < R^{\mathrm{R}}/[1 - (1 - m_{\mathrm{ct}})R^{\mathrm{R}}]$.