Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 2017–2021, April 1983
Genetics

Paradox of the evolution of communication and of social interactivity

(cooperativeness / altruism / social interaction / evolutionary bottleneck)

L. L. Cavalli-Sforza* and M. W. Feldman†

Departments of *Genetics and of †Biological Sciences, Stanford University, Stanford, California 94305

Contributed by L. L. Cavalli-Sforza, January 3, 1983

ABSTRACT Communication between individuals of a species is likely to increase the capacity to acquire skills useful for survival and propagation and thus may confer important selective advantages. Since interaction occurs between two or more individuals, the selective process is frequency dependent, and the analysis shows that communication cannot initially increase at a reasonable rate when it is limited to random unrelated individuals, so that it is likely to abort for stochastic reasons. However, this bottleneck is removed if the communication process takes place in the nuclear family or among close relatives or if aggregation of communicators occurs because of assortative mating or meeting. Use of the individual conditional fitnesses we have introduced earlier permits an exact analysis. We show that, in general, the initial rate of increase can be geometric if and only if, in the class of selective models considered, the conditional probability of a communicator interacting with another contains a positive constant term. In our discussion of communication, cost factors for the act of communication have been omitted. However, the model has been generalized to include cooperativeness, and also altruism, or competition, by introducing costs. There is a close relationship among these situations, and the same considerations about the initial bottleneck and its resolution also extend to them. The models given here are for haploids but they extend to diploids and the conclusions are similar.

We have studied some theoretical aspects of the problem of evolution of communication between two or more individuals, the key ingredient of cultural behavior. Communication is taken to include the range extending from the simplest examples—for instance, the recognition of a parent in an imprinting situation or the capacity to learn a skill by imitation—to the most complex ones involving communication by language. When there is genetic variation in the capacity to acquire or to communicate a specific ability that is subject to natural selection, it seems intuitively likely that there will be evolution toward more efficient communication if the skills thus learned have positive selective value (and vice versa if negative). In the course of the theoretical treatment, it was found that there is a bottleneck in the evolutionary process of communication at its very beginning.

Not surprisingly, there are similarities between communication, cooperativeness, sociality, altruism. Axelrod and Hamilton (1) have explicitly drawn attention to the existence of a dilemma, whereby a genetic trait determining cooperativeness that is initially rare (as all new mutants are) cannot increase in a population. They base their argument on the approach by game theory. Using this same approach, Eshel and Cavalli-Sforza (2) have shown that it is possible for cooperativeness to increase if there is any tendency for cooperative individuals to aggregate, because of assortative mating or meeting (passively or actively sought). Boorman and Levitt (3, 4) introduced a diploid genetic model of fitness gain to social interactions in which contacts between social individuals resulted in a fitness gain to both. A fitness loss was experienced by social individuals who contacted asocial types and no decrement resulted from contacts between asocial individuals. This obviously results in a frequency-dependent type of underdominance that has an intermediate frequency threshold above which sociality will spread. Our problem concerns the difficulty of the early stages of the evolution in which communication is rare, a problem not really addressed in the demonstration by Boorman and Levitt of a critical intermediate threshold.

In this paper, we show how the dilemma can be solved in a more general way. We will use the method of individual conditional fitness that we have introduced for the treatment of kinship (5), which we suggest can be generalized to a great variety of interactions between individuals.

In principle, communication will always involve a reciprocal interaction between a communicator and a recipient. More than one individual of either category may be involved, but we will limit our consideration to interaction in pairs. More general treatments will consider separately the abilities to communicate, to receive the communication, and to actually acquire the skills that are communicated. In simpler cases, variation may exist for only one of these, but there may well be variation for all. We will not consider effects of variation for the third capacity here but will elaborate models in which there is joint variation in the first two abilities. Separate variation in the communicating (teaching) and receiving (understanding) abilities will be considered elsewhere.

The models developed here show that cultural interactions within the nuclear family or among close relatives can have considerable importance in fostering the evolution of communication. They also indicate that assortative mating and meeting may play a substantial role. In addition, we will show that these models can also be viewed as special cases of a general model of evolution of cooperation and altruism, to which results on communication are easily generalized. The treatment will mainly be limited to haploids with sexual recombination (or equivalent transmission schemes). Only simple cases of diploids will be treated in connection with earlier results (5).

Communication in haploids

We first consider a model of haploid transmission with random mating. As in our earlier work on "altruism" (5), fitness is regarded as a property of the individual. Communication *per se* probably has no direct effect on fitness but, if it leads to the acquisition of information that can increase individual fitness,

---

Abbreviation: CK, conditional kinship (matrix).

then its indirect selective effect may be important. Of course, for an effect on the fitness of an individual to be realized, other individuals capable of communication must be present; the passage of information requires a minimum of two individuals.

We will assume for simplicity that (*i*) communication is reciprocal; i.e., individuals who are communicators can also receive information; (*ii*) all individuals can potentially learn a skill for survival if they are communicators; and (*iii*) all adult communicators have acquired the skill that confers increased fitness.

The individual fitness of a communicator (type *A*) who has been in contact with other communicators will be $1 + \beta$, relative to fitness of 1 for noncommunicators (type *a*). At generation *t*, the frequency of communicators is u_t ($v_t = 1 - u_t$), and at birth the offspring frequency of *A* is again u_t. However, by communicating with other *A* individuals, a fitness increment proportional to their frequency accrues to the communicators, whose fitness is therefore

$$1 + s = 1 + \beta u_t, \qquad [1]$$

where a constant β measures effective communication. Thus, after communication, the new frequency, u_{t+1}, is

$$u_{t+1} = \frac{u_t(1 + \beta u_t)}{1 + \beta u_t^2}. \qquad [2]$$

The denominator is obtained from the sum of communicators $u_t(1 + \beta u_t)$ and noncommunicators, v_t, after selection. We suppress the time subscript and write *u* for u_t, u' for u_{t+1}. Naturally $u' > u$ if $\beta > 0$, so that *u* increases monotonically and communication will spread in the population. But the key evolutionary problem is at the beginning—i.e., at the time when there are very few communicators and therefore *u* is close to zero. The rate of increase is

$$\frac{u'}{u} = \frac{1 + \beta u}{1 + \beta u^2} \qquad [3]$$

and, taking the limit of this ratio for $u = 0$, the initial rate is one, so that the trait for communication cannot increase at a geometric rate. At most, the increase is very slow and this is probably the explicit formulation of the evolutionary paradox already noted by Axelrod and Hamilton for cooperativeness.

In the case of communication, we have a simple proposal to solve the paradox. Communication (e.g., language in man) is usually learned first in the family—i.e., by the child from the parents (and sibs and possibly other close relatives). To model this, we follow our earlier approach (5) of conditional kinship matrices (CK), which allow us to compute the probability that a relative (of a given degree) of an individual with a given genotype has a specified genotype. This matrix, in the case of haploids and the child–parent relationship, is most simply calculated from the frequencies of types of matings and their progeny (random mating assumed), as shown in Table 1.

The probability that an A_i child is born from an A_j parent (A_i or A_j can be here *A* or *a*) independently of parental sex, is equal to

$$f_{A_j|A_i} = P(\text{parent } A_j|\text{child } A_i) = \frac{\sum^{k} P_k \cdot P(A_j|k) \cdot P(\text{child } A_i|k)}{\sum^{k} P_k \cdot P(\text{child } A_i|k)}, \qquad [4]$$

where P_k is the relative frequency of the *k*th mating, $P(A_j|k)$ is the relative frequency of A_j parents among the parents in the *k*th mating [$P(A_j|k) = 1, 1/2$, or 0]; and $P(\text{child } A_i|k)$ is the probability of an A_i child in the *k*th mating. The sums are extended to all matings. Then, the CK matrix for parent/child haploids is

Table 1. Probability that progeny of a mating will have a given genotype

		Probability of progeny	
Mating	Probability	*A*	*a*
$A \times A$	u^2	1	0
$A \times a$	$2uv$	$^1/_2$	$^1/_2$
$a \times a$	v^2	0	1

$$\begin{array}{c|cc} & \multicolumn{2}{c}{\text{Parent}} \\ \text{Child} & A & a \\ A & u + v/2 & v/2 \\ a & u/2 & u/2 + v \end{array} \qquad [5]$$

The fitness of an *A* individual does not depend only on the frequency *u* of the trait in the parental generation but also on the probability that the individual has learned from his parents. The fitness of the *A* type will be

$$1 + \beta_0 u + \beta_1 (u + v/2), \qquad [6]$$

where the $\beta_0 u$ component derives from adult contact with the general adult population (as in Eq. 1) and $\beta_1(u + v/2)$ derives from contact with the parents, at least one of whom will be *A*. With fitness (Eq. 6), we have

$$u' = \frac{u(1 + \beta_0 u + \beta_1 (u + v/2)}{1 + u(\beta_0 u + \beta_1 (u + v/2)} \qquad [7]$$

and the value of the rate of initial increase is

$$\lim_{u \to 0} \frac{u'}{u} = 1 + \beta_1/2, \qquad [8]$$

which is >1 if $\beta_1 > 0$, so that the contribution from parental teaching allows the evolutionary process to commence. Later, the contribution from other members of the parental generation, which we have called "oblique" in the context of cultural transmission (6), will contribute to make the process faster. At the end of the process, the rate ($\lim v'/v$, $v \to 0$) for fixation of communicators is $1/(1 + \beta_0 + \beta_1)$ and here, as during the rest of the process for $u > 0$, the oblique contribution is important.

Sib–sib communication requires another CK matrix. The general term of this matrix is

$$f'_{A_j|A_i} = P(\text{sib } A_j|\text{sib } A_i) = \frac{\sum^{k} P_k \cdot P(\text{child } A_i|k) \cdot P(\text{child } A_j|k)}{\sum^{k} P_k \cdot P(\text{child } A_i|k)}. \qquad [9]$$

In the case of haploids with random mating, the sib/sib CK matrix is identical to the parent–child matrix, and the evolutionary conclusion is therefore also identical. Adding to the oblique term the contributions from parent–child (with β_1) and sib–sib (with β_2), the selection coefficient is

$$s = \beta_0 u + \beta_1 f_{A|A} + \beta_2 f'_{A|A} \qquad [10]$$

and this allows a faster initial increase than with Eq. 6. Note that the addition of contributions to fitness from different sources as above can be the basis for an explicit definition of inclusive fitness, which is so far missing (see, however, ref. 7).

Assortative mating and meeting

Communicators may mate nonrandomly, and in fact it is likely that there will be positive assortment, so that matings A × A will be more frequent than others. We will use three schemes of assortative mating (Table 2): the first is the usual one initially proposed by Wright (8) and gives rise to expectations for the three mating types shown for model 1. In model 2, we assume that only communicators choose; they choose other communicators with probability m, and subsequently the residual population, made of $(1 - m)\,u\,A$ and $v\,a$ individuals (total, $1 - mu$) mates randomly. This model is appropriate when sexes are in equal numbers and mating of all possible pairs takes place, so that there is no selection due to mating according to the definition of strict assortative mating (9), which also applies for model 1. In model 3, it is assumed that one sex, which is present in smaller numbers (e.g., females) chooses the other sex, which is in virtually infinite numbers. A proportion m of A females chooses A males, and the rest $(1 - m)\,u\,A$, $v\,a$ of the females choose randomly from $u\,A$, $v\,a$ males. Here there is sexual selection because u' increases constantly. In fact, taking u' as the frequency of $A \times A$ matings + 1/2 that of $A \times a$ matings, $u' = u$ in the first two models, while, in the third,

$$u' = u + muv/2, \qquad [11]$$

so that u increases continuously. At $u = 0$, its rate of increase is $1 + m/2$ and, at $v = 0$, v'/v is $1 - m/2$. Thus, evolution proceeds from any $u > 0$ to a stable equilibrium at $u = 1$, even in the absence of any other type of selection.

The one-generation transformation in the first two models is

$$u' = \frac{u(1 + \beta f_{A|A})}{1 + \beta u f_{A|A}}, \qquad [12]$$

where $f_{A|A}$, the probability that the parent of a child A is A; i.e., the first element of the CK matrix is

$$\text{Model 1: } f_{A|A} = m + (1 - m)(u + v/2)$$

$$\text{Model 2: } f_{A|A} = \frac{u + (v/2) + m\,(1 - 3\,u)/2}{1 - mu}. \qquad [13]$$

The same results are obtained for the sib/sib matrix.

The initial increase of u occurs at rate

$$\lim_{u \to 0} \frac{u'}{u} = 1 + \beta(1 + m)/2 \qquad [14]$$

in both models. Hence, if assortative mating is complete ($m = 1$), the rate of increase is $1 + \beta$ and, if $m = 0$, it is that previously seen for parent–child or sib–sib.

For model 3, there is selection due to mating and hence we must consider both stages of selection. Let u^s be the gene frequency after selective mating. Then, from Eq. 11,

$$u^s = u + muv/2,$$

and

$$u' = \frac{(1 + \beta f_{A|A})u^s}{1 + \beta u^s f_{A|A}} \qquad [15]$$

is the frequency in the offspring. Both u and u' are estimated after selection due to communication and before selection due to mating.

Table 2. Assortative mating of communicators: three models

Mating	Probability of mating Model 1	Model 2	Model 3
$A \times A$	$mu + (1 - m)u^2$	$mu + \frac{(1 - m)^2u^2}{1 - mu}$	$mu + (1 - m)u^2$
$A \times a$	$2(1 - m)uv$	$2uv(1 - m)/(1 - mu)$	$(2 - m)uv$
$a \times a$	$mv + (1 - m)v^2$	$v^2/(1 - mu)$	v^2

Here,

$$f_{A|A} = \frac{u + (v/2) + 3\,mv/4}{1 + mv/2}, \qquad [16]$$

and

$$\lim_{u \to 0} \frac{u'}{u} = (1 + m/2)[1 + \beta\,(1 + 3m/2)/(2 + m)]; \quad u = 0 \text{ is unstable,}$$

$$\lim_{v \to 0} \frac{v'}{v} = (1 - m/2)/(1 + \beta); \qquad u = 1 \text{ is stable.}$$

In all these cases, assortative mating speeds up the initial increase of a trait allowing communication, provided there is communication in the family (parent–child or sib–sib and also more remote relatives).

As shown in ref. 2, assortative mating (or meeting) *per se* can bring about the initial increase of cooperativeness. This is predicated on the assumption that individuals form pairs in which (A,A) pairs are more frequent than under random union. Communication takes place *after* pair formation, and selection occurs afterward on these pairs. If the pairs are mates, communication between the members of a pair gives greater fitness (perhaps in the form of higher fertility) to the $A \times A$ pair as compared with $A \times a$ or $a \times a$. In this case, the use of model 1 of assortative mating, which can be extended to meeting, gives the results shown in Table 3.

Hence, after selection,

$$u' = \frac{(1 + \beta)[mu + (1 - m)u^2] + uv(1 - m)}{1 + \beta[mu + (1 - m)u^2]} = \frac{u\{1 + \beta[m + (1 - m)u]\}}{1 + \beta u[m + (1 - m)u]} \qquad [17]$$

and

$$\lim_{u \to 0} \frac{u'}{u} = 1 + \beta m.$$

The same result is obtained by using model 2. For model 3, the rate of initial increase is $1 + m(1 + 2\beta)/2$.

The same formulas are valid if, instead of mating pairs, we think of cooperative interaction between pairs of unrelated individuals, leading to increased fitness of members of the A,A pair. In Eq. 17, $m + (1 - m)u$ takes the place of the conditional probability that an A individual meets another A individual, because this is the probability of formation of an (A,A) pair divid-

Table 3. Fitness of pair: model 1 of assortative mating

Pair	Probability	Fitness of pair
A,A	$mu + (1 - m)u^2$	$1 + \beta$
A,a	$2uv(1 - m)$	1
a,a	$mv + (1 - m)v^2$	1

ed by the probability of an A individual: $[mu + (1 - m)u^2]/u$.

Thus, assortative mating, or assortative meeting of individuals in pairs, can overcome the impasse of the lack of increase of communication near zero frequency, provided selection takes place after the interaction of the members of the pair.

The condition common to all cases we have seen, in which the initial increase of communication takes place at a geometric rate, is that the conditional probability of contact of the communicator with another communicator (be it a parent, a sib, a spouse, or an unrelated individual) contains a constant positive term, independent of u. In other words, $\lim_{u \rightarrow 0} f_{A|A} > 0$. This condition is clearly sufficient for initial increase to take place.

Within the class of models specified by Eq. **12**, and more generally Eq. **15**, it is clear that, for geometric increase (i.e., $\lim_{u \rightarrow 0} u'/u > 1$) it is necessary that $\lim_{u \rightarrow 0} f_{A|A}$ be positive.

Relations with cooperativeness, altruism, and the game theory approach

The model of communication is formally very simple. But the conclusions are easily generalized to situations that include communication, cooperation, altruism, kin selection, etc. In all these cases, individuals interact in pairs and the interactions have an effect on fitness. We have assumed that communication has no cost but only equal benefit to both people who communicate. A slightly more complex model is one in which the abilities to communicate and to receive information are to some extent independent, a being more able to receive than to give information, so that, in pairs $A \times a$, a has an advantage with respect to A. This model is indistinguishable from the simple model of cooperation shown in Table 4. If costs are introduced, we have a model indistinguishable from the usual ones of "altruism" summarized in Table 4. All formulas written for communication can be adapted by the inclusion of a fitness loss γ, and they are also valid for the case of altruism (haploid). The diploid case has already been given in this more general form and can be reduced to the case of communication by setting $\gamma = 0$.

In all conditions for initial increase, incorporation of the cost term γ will simply result in subtraction of γ from $\lim_{u \rightarrow 0} u'/u$, so that, for instance for altruism between members of pairs in assortative mating or meeting, the condition for initial increase would become

$$\gamma/\beta < m \qquad [18]$$

and for parent–child or sib–sib altruism (with assortative mating of the model 1 or 2 type), the condition is

$$\gamma/\beta < (1 + m)/2. \qquad [19]$$

If the more general notation given for the game theory analogue in Table 4 is used,

$$V_{12} - V_{22} = (V_{12} - V_{11}) \lim_{u \rightarrow 0} f_{A|A} + (V_{21} - V_{22}) \lim_{u \rightarrow 0} f_{a|A} \qquad [20]$$

is the condition for initial increase of A, with the sign of the inequality being reversed for the condition for initial increase of a when it is rare, in agreement with ref. 2.

Using the CK matrices developed for diploids as an aid to the discussion of altruism (5, 9), one can see that the condition for initial increase is $\beta \lim u \rightarrow 0 f_{12} > \gamma$, where f_{12} is the probability that a heterozygote child has an altruistic relative of given degree (i.e., a heterozygous relative for $u \rightarrow 0$). In the formulation of Cavalli-Sforza and Feldman (5) for random mating—e.g., $f_{12} = \beta/2$ for parent–child or sib–sib— and with additive fitnesses, the result is identical to that for haploids. The case of communication is easily obtained by setting $\gamma = 0$, so that, if β and $\lim f_{12} > 0$, communicators will always increase.

Discussion

Communication is likely to be of considerable selective value if it can provide individuals with information that is already available to other individuals of the same population and increases the viability of progeny. In principle, this kind of information can also be obtained from individuals of other species, such as when a potential prey is alerted to the presence of a predator by the behavior of individuals from other species. Our analysis, however, applies only to the acquisition of information from one or more other individuals of the same species. Language is probably the most developed form of communication in existence, and its importance in the acquisition of skills need not be elaborated upon.

At the beginning of the selective process, there may only be one individual, say a mutant, possessing higher capacity for communication. As communication requires interaction between two individuals, one (or few) communicators may be at no special advantage if there is no one to communicate with. Therefore, the increase in fitness to communication is strictly frequency dependent, and the initial period of evolution may be critical. In fact, we have seen that, without special stipulations, the initial rate of increase of communication is such that the evolutionary process is not capable of geometric increase. It will therefore be so slow that it will very likely be aborted by chance events. But if, for one reason or another, the frequency of communicators at the beginning is not neglible, then even the first phase can be rapid, as in an ordinary process of selection with constant coefficients.

It is fairly obvious that, in most instances, communication is established first in the nuclear family between parents and children or between sibs. Here, the local concentration of communicators can be high even if only one parent is a commu-

Table 4. Fitness values in models for various types of social interaction

			Fitness of type in pair			
	Phenotypes (genotypes)			*A,a* pair		
Interaction	*A*	*a*	*A,A* pair	*A*	*a*	*a,a* pair
Communication	Communicator and recipient	Unable to communicate	$1 + \beta$	1	1	1
Cooperation	Cooperative (toward anybody)	Uncooperative	$1 + \beta$	1	$1 + \beta$	1
Altruism	Altruist	Selfish	$1 + \beta - \gamma$	$1 - \gamma$	$1 + \beta$	1
Game theory analogue*			$1 + V_{11}$	$1 + V_{12}$	$1 + V_{21}$	$1 + V_{22}$

* See ref. 2.

nicator. Thus, the advantage of communication can be established immediately and, even if there is only one communicator in the whole population, the first nucleus of communication will be the family this individual begets. Ignoring stochastic factors that affect the probability of every initial increase, the evolution of communication can thus begin promptly.

Here, we have presented a treatment of the evolutionary increase of communication between parent–child or sib–sib and with unrelated invididuals (oblique transmission). This is made possible by the use of individual fitnesses, conditional on the frequency of communicators among relatives; the same treatment can be extended to interaction with nonrelatives, as in the three models of assortative mating (or meeting) also considered. In fact, we show that the determinant of the initial increase is the constant term in the conditional probability that the communicator interacts with another communicator. In the class of models here considered, the necessary and sufficient condition for geometric initial rate of increase of communication is that the constant term of this conditional probability is positive.

While it is possible that there is no fitness cost, only benefit, to being a communicator, it is easy to introduce fitness costs. The more general models thus generated show the close relationships existing between communication, cooperativeness, altruism, and in general social interactions, so that one can with a single model summarize all these situations.

Thus, our analysis shows that there is indeed a bottleneck at the beginning of the evolutionary process, when a mutant with increased communication, cooperativeness, or altruism (not directed necessarily to close relatives) is first introduced into a population, so that, under interaction between random unrelated individuals, the evolutionary process will be so slow that the increase of communicators, etc., will in practice be negligible. However, if interaction begins in the nuclear family (and this hypothesis can be extended to close relatives), or if positive assortative mating is present; or if individuals meet assortatively, the bottleneck is removed and the initial increase of the interaction takes place at a geometric rate. It will be of interest to evaluate the roles of these or similar conditions of relationship or aggregation in empirical situations, for one can predict that, in their absence, evolution is unlikely to begin in spite of the existence of initially rare variants capable of increased interaction.

Our models are for haploids, although some of these conclusions have been extended to diploids by using the same procedure. Moreover, we do not distinguish the capacities to acquire and to give information, and we assume that both are increased in the new genetic type A, called "communicator" in comparison with the older type a, called noncommunicator. A distinction between the two will be given in another paper.

1. Axelrod, R. & Hamilton, W. D. (1981) *Science* **211**, 1390–1396.
2. Eshel, I. & Cavalli-Sforza, L. L. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1331–1335.
3. Boorman, S. A. & Levitt, P. R. (1973) *Theor. Pop. Biol.* **4**, 85–128.
4. Boorman, S. A. & Levitt, P. R. (1980) *The Genetics of Altruism* (Academic, New York).
5. Cavalli-Sforza, L. L. & Feldman, M. W. (1978) *Theor. Pop. Biol.* **15**, 276–307.
6. Cavalli-Sforza, L. L. & Feldman, M. W. (1981) *Cultural Transmission and Evolution: A Quantitative Approach* (Princeton Univ. Press, Princeton, NJ).
7. Cohen, D. & Eshel, I. (1976) *Theor. Pop. Biol.* **10**, 276–302.
8. Wright, S. (1921) *Genetics* **6**, 144–161.
9. Lewontin, R. C., Kirk, D. & Crow, J. (1968) *Eugen. Quart.* **15**, 141–143.