

✓ 2-1391

WORKSHOP IN POLITICAL THEORY  
AND POLITICAL ANALYSIS  
550 NORTH PARK  
INDIANA UNIVERSITY  
BLOOMINGTON, INDIANA 47405

Excerpt  
A 5016  
A 202

Ulrich O. Mueller\*, Benzion Chanowitz\*\*, Ellen Langer\*\*

The Rationality of the Free Ride

Paper Prepared for the 1983 Annual Meeting of the Public Choice Society in  
Savannah, Georgia, March 24 - 26, 1983

\* Massachusetts Institute of Technology / Freie Universitaet Berlin

\*\* Harvard University

In this paper I will argue\* the following points:

1) Prescriptive and descriptive solutions to the free rider problem are insufficient, when they either demand the establishment of institutions or point to the existence of institutions which are able to prevent people from taking a free ride, but fail to explain how these institutions can actually develop during biological or cultural evolution .

2) That a certain behavior is good for the group or the society does not necessarily mean that it has a chance to evolve and to become stabilized if it is potentially harmful to the individual who displays this behavior. This is possible only under certain conditions the study of which is an important issue of modern theoretical population biology. Some of the relevant findings will be described.

3) These findings have an enormous impact on the further lines of development of theoretical as well as empirical research in the public goods/free rider field. A conceptual framework for theoretical research as well as the design of an experimental study under way will be presented.

\*The theoretical argumentation in this paper reflects the viewpoints of the first author. The experiment has been designed and is being conducted at the Department of Psychology and Social Relations, Harvard University, in cooperation with the other two authors.

The free rider problem is usually stated in such a way that the achievement of a Pareto optimal allocation of resources in the acquisition of public goods is fundamentally incompatible with purely individual incentives. This fact is regarded as a problem because it is obviously bad for society as a whole and the single producer/consumer. Solutions have to be found which will make people reveal their true preferences, refrain from taking free rides, and remain faithful members even of large groups in order to make a optimal or at least better supply of public goods possible.

The problems as stated in this way has been attacked on theoretical as well as on empirical levels:

Theoretical approaches developed sophisticated sets of rules which should make it advantageous for each consumer to state his true preferences - the probably best known of these models are the ones by CLARKE (1971), TIDEMAN/TULLOCK (1976) and GROVES/LEDYARD (1977). These and other models have in common that they do not make as the basis for the pricing mechanism the total number of units of the public good divided by the number of people benefitted by this public good, but instead the increment or the decrement a consumer wants to see added or subtracted from the public good in question.

It is clear that these rules and mechanisms do not just fall down from heaven. What these models are actually looking for is a "class of optimal government rules " (GROVES/LEDYARD 1977; p.791), which are not likely to emerge spontaneously in elementary encounters between people because otherwise there would be no need for inventing them with so much intellectual sophistication. So we have not only to ask whether these rules can make people refrain from taking free rides - with all the effects this would have - but also how these rules can be established - a point which the mentioned authors remain

remarkably silent.

On the empirical level one has to deal with the undeniable fact that there obviously exist large organizations which provide certain public goods to members as well as non-members, even beyond that point in size where a single actor's contribution can make a tangible difference the formation of groups is possible. So the task is seen as identifying those forces which make that possible. All given answers fall into three categories:

*By Product*

- a) A clever combination of public goods with private goods in a way that making contributions to a public good can be rewarded with privileged access to certain private goods, and taking free rides can be penalized by withdrawal of this access.
- b) Group norms which encourage fidelity towards the group, equal share of the burdens for public goods which the group has decided to acquire, and denounce free riding.
- c) Sheer force, applied by some centralized power within a group, or a whole society.

It is not difficult to see the common features not only in these three categories, but in both the empirical and the theoretical field. It is the search for institutions which ensure that people don't behave in the way the Samuelson/Olson theorem states. But what are institutions other than stable behavior patterns of groups of people? How satisfactory is an answer to the free rider problem, as defined in the beginning, which actually tells that one should (the theoretical level), or one has already (the empirical level) established a behavior pattern which enables the supply of public goods and the formation of large groups? Obviously not very much.

A solution to the free rider problem requires an explanation of how such an anti-free-ride behavior pattern can be stably established within a group. The Samuelson/Olson theorem clearly predicts that all these patterns are unstable. If it is advantageous for everyone to take a free ride, and therefore risky for everyone not to take one, why then stick to certain rules, as sophisticated as they might be, if there are no individual incentives for enforcing them?

So we are asking for a solution in evolutionary terms: how can institutions/group behavior patterns evolve which can prevent its members from taking free rides? Not before we are able to answer this question, can a description as well as a prescription of such an institution be accepted as a satisfactory solution to the free rider problem.

We have been trained to assume that by nature as well as by culture our motivation patterns are such that purely individual incentives are not the only ones, and that caring for relatives, friends, the community, etc. is built in our mind as a equivalently important goal.

Putting it in this way is a step in the right direction since it indicates that even from the best institutions we can only expect that they will make a non free rider behavior the dominant mode, but no human institution can be designed to be absolutely free rider proof. Later on we will see more clearly why this is so.

But otherwise this answer must inevitably lead to another question: how can non individual incentives obtain such an important role in our motivational endowment? The vast majority of mammals have no or only a very rudimentary

In fact, the search for models which can explain the development of social behavior with a non purely individual incentive structure is the most important branch in recent population biology theory. We have to take a close look into these models.

The well known label of these recent developments is the evolution of altruism. But this title certainly covers too much.

We are looking for models which can explain how a behavior pattern or trait which benefits others more than its bearer, nevertheless can evolve and spread out under the conditions of Darwinian evolution. This implies that - unlike the usual classification - the development of "reciprocal altruism" (prototypical BOORMAN/LEVITT 1973) as well as of "kinship altruism" (prototypical HAMILTON 1964a, 1964b) should be excluded from the realm of true altruism.

The first should correctly be named the development of cooperativeness; cooperativeness is only occasionally altruism if the initial cooperation by the first actor is not responded to by an equally cooperative move by the other actor, but by a move which harms the first actor.

As soon as cooperative behavior on a reciprocal basis is firmly established, it clearly benefits all participants, and the only remaining problem is then the deterrence of defections by effective methods of early detection and retaliation as well as rewarding of faithfulness. It is obvious that as far as public goods are concerned such a model cannot explain the development of a free ride avoidance behavior.

The key notion of kinship altruism is inclusive fitness, which means that for a strategy of maximizing the fitness of one's own genes it can be advantageous to forego own fecundity or even survival chances in favor of a relative ( child, parent, sibling etc. ), if  $k > 1/r$ , where  $r$  is the coefficient of

social organization, and the human history is a history of the gradual development of more and more complex social institutions with their according motivational internalizations.

So the free rider problem has to be properly reformulated: Given the impeccable logic of the SAMUELSON/OLSON theorem, how can a non free rider behavior evolve and get stabilized into group behavior patterns as described? This question rises on the descriptive level in the same way as on the prescriptive level. Both levels are thus being brought together in the evolutionary perspective.

## II

The usual answer to this question, whether given explicitly or implicitly is that since it is good for society/the human species, a general behavior of avoiding the exploitation of free rider options must have developed, be it on the level of the genetic endowment or of the basic cultural patterns. The argument goes that surely individual incentives are not favored by biological as well as cultural evolution, that our social behavior (as well as the one of "higher" animals) is governed by a mixture of individual and collective incentives.

In principal, however, one should expect that Darwinian Evolution precisely does not favor the development of a non purely individual incentive structure in the individual actors, because the more members of the group or the whole society do not cheat, the higher the reward = selection pressure for cheating obviously becomes.

relatedness between donor and recipient of this fitness transfer, and  $k$  the ratio of gain of fitness to the recipient to the loss of fitness of the donor. Clearly the object of selection is here not the individual but the kin, and again we should not call this form of fitness transfer altruism. Furthermore it is hard to see how the notion of inclusive fitness could be made applicable for models of the development of social institutions.

So we have to start our considerations at the remaining category: group selection of altruism. The basic idea here is (expressed for the haploid, monolocal case) that, if a population is scattered into clearly distinct groups or demes, an allele 'a', which is more beneficial for the average fitness of the group members than the 'A' allele, but which yields less individual fitness to its bearer, nevertheless can outgrow the latter - if certain conditions are fulfilled.

The refinement of the model and the specification of its parameters has been, ever since the pioneer work by HALDANE (1932) and WYNNE-EDWARDS (1963), the subject of an increasingly sophisticated branch of theoretical population biology. Recent contributions - even more exciting - have furthermore demonstrated the possibility of the evolution of altruism without two basic features of classical group selection models, namely differential extinction rates for the groups as a whole, and random genetic drift within these groups with consequential fixation of one allele.

These new approaches, however, are still using the concept of an at least temporary separation of large populations into small scattered groups ( demes, founder populations ) which provide different internal selection milieus. But even this assumption can be replaced, so that we are now able to model the development of altruism under conditions of Darwinian evolution in large



homogeneous populations.

Before we start with the classical group selection model, a word about the applicability of results of theoretical population genetics to problems of economics or other social sciences.

The Darwinian theory of evolution deals with long run changes in phenotypes due to the forces of natural selection. It is, however, impossible to understand and predict the effects of natural selection on a given trait without some idea of the mechanisms by which this trait is transmitted from one generation to the next.

The problem is only that the genetic basis for evolutionary changes is in most cases not known. Moreover, there is ample evidence that most traits have a multilocal genetic basis which makes an analysis with the methods of classical population genetics soon hopelessly complex.

So the main line of present research goes into an different direction, as advocated by ESHEL (1982, p.214): " An alternative approach to the study of this sort of traits attempts to avoid complications stemming from the specific nature of one genetic structure or another. Instead, intuitively understood criteria of phenotype optimization are suggested, with the basic assumption that . . . the basic Darwinian relation between adaption and natural selection must lead to some sort of local optimization, at least as a workable approximation. Thus instead of dealing with many technical unmeasurable, and presumably insignificant details, it is preferable to ignore them in order to obtain simple qualitative results." Such models of phenotypic optimization can be developed and tested independently from assumptions about the specific nature of the transmission mechnism, whether it is the reproduction of genetic material, or some cultural learning process.

State-of-the-art models of the group selection theory can be found in WILSON (1975, 1979), GILPIN (1975) or BOORMAN/LEVITT (1980). I will here refer to the BOORMAN/LEVITT model. GILPIN assumes permanently separated demes with migration, which makes his approach more complicated; WILSON, on the other hand, mixes group selection with kin selection ideas (MAYNARD-SMITH 1976). Given is a infinite population, divided into demes except for recolonization, a biallelic locus ('A','a'), where 'a' is favoured by group selection, and 'A' by individual ( "Mendelian" ) selection. Phenotyp A ,the altruist, is borne by all 'aa's, and the nonaltruist N by all 'AA's. It is assumed that the individual selection pressure is that strong that any deme population reaching carrying capacity in one period of time will have lost heterozygosity at the ('A','a') locus, so we will never have to deal with a phenotype borne by Aa. As usual, time is periodized in this model into discrete periods. Within each period the following sequence of events occurs: 1) extinction, 2) growth, 3) recolonization. At the beginning of each period, one can distinguish five classes of demes, two in the class of completely occupied/carrying capacity demes, and three in the class of founder populations, i.e. just recently recolonized demes. The model works with the simplest form of all conceivable founder populations, comprising one female and one male. So we have:

- I) N/N
- II) A/N
- III) A/A

and the carrying capacity demes

- IV) all homozygous for 'AA's

V) all homozygous for 'aa's

Further assumptions are

- growth of newly founded populations is that fast, that a population reaches carrying capacity ( $N = K$ ) within one time period, if it escapes extinction at the beginning of this period.
- if a deme is emptied by extinction at the beginning of a period, it remains empty for the whole period and is recolonized at the end of it by a random sample from the gene pool of carrying capacity demes.
- the founder populations are subject to differential extinction probabilities  $0 < E_i < 1$  ( $i = 1, 2, 3$ ), where group selection favours the 'a' gene:  $E_1 > E_2 > E_3$ .
- in addition to this differential extinction acting on founder populations there exists a uniform extinction capacity  $0 < E < 1$  which acts on carrying capacity populations at the beginning of each time period. In the event of extinction, recolonization takes place in the same way as described extinct founder populations.
- all parameters specifying the individual selection are expressed in the parameter 'u' such that mixed founder populations of class II (N/A) grow to a carrying capacity population of class V with probability u and to class IV with probability (1 - u), while founder populations of class I or III grow to class IV or V respectively. From the definition of altruism it is required that  $0 < 'u' < (1 - u) < 1$ , or simply  $'u' < 1/2$ .

The dynamics of this model can be analyzed with the following system of difference equations

$$\vec{\gamma}(n+1) = \vec{\gamma}(n) \Gamma \quad (1)$$

with the I.C.  $\vec{\gamma}^0 = [\gamma_1(0), \gamma_2(0), \gamma_3(0), \gamma_4(0), \gamma_5(0)]$  at time  $n = 0$

$$0 < \gamma_i < 1 \quad \text{and} \quad \sum \gamma_i = 1 \quad (i = 1, 2, 3, 4, 5) \quad \text{of the } i\text{th class.}$$

and  $\Gamma$  the transition matrix according to the listed assumptions.

The system has two obvious equilibrium points, the one when only N - phenotypes exist, the other one when only A - phenotypes exist, and the extinction rates fulfil the conditions

$$\vec{\gamma}^A = \frac{1}{1 + E - E_1} (E, 0, 0, (1 - E_1), 0) \quad (2)$$

which represents an 'A' - fixation, or

$$\vec{\gamma}^a = \frac{1}{1 + E - E_3} (0, 0, E, 0, (1 - E_3)) \quad (3)$$

for the 'a' fixation.

More interesting for our argumentation; of course, is the analysis of an internal equilibrium point, i.e. when all classes are occupied, which then

entitles us to speak about a polymorphic equilibrium.

BOORMAN/LEVITT can show that a necessary and sufficient condition for this is either

$$\frac{E_3 + 1 - 2E_2}{2(1 - E_2)} < U < \frac{1 - E_1}{2(1 - E_2)} \quad (4)$$

or

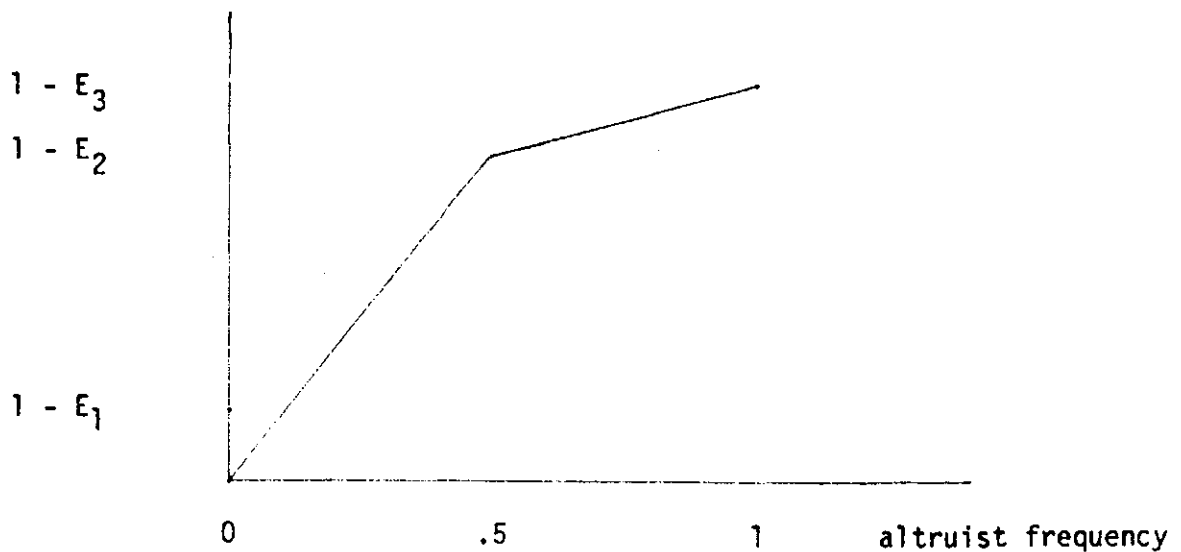
$$\frac{1 - E_1}{2(1 - E_2)} < U < \frac{E_3 + 1 - 2E_2}{2(1 - E_2)} \quad (5)$$

Both conditions are obviously mutually exclusive.

By considering two distinct cases, depending on whether

$$(1 - E_2) \begin{cases} \leq \\ > \end{cases} (2 - E_1 - E_3)/2 \quad (6)$$

it can be show that only in the convex case  $(1 - E_2) > (2 - E_1 - E_3)$  a stable polymorphic equilibrium is possible.



while in the concave case  $(1 - E_2) < (2 - E_1 - E_3)$  any internal equilibrium will be unstable. This means that only condition (5) will allow for an stable polymorphic equilibrium. BOORMAN/LEVITT can furthermore show that an internal equilibrium if it exists at all, will be globally stable, which means that the edge points which represent pure 'a' or pure 'A' fixation will then be unstable.

This is the first of the two important results which we can obtain from the analysis of the BOORMAN/LEAVITT model. A globally stable polymorphism of altruists/nonaltruists is possible if and only if the differential extinction function  $E_i = f(i)$ ,  $i = 1, 2, 3$ , is strictly convex.

The second result can be obtained by a different approach to the problem of stable polymorphism. BOORMAN/LEVITT obtain their inequalities (4) and (5) by an analysis of the system  $\vec{y}(n+1) = \vec{y}(n) \mathbf{M}$ . We will demonstrate an alternative path to alternative (5), the necessary and sufficient condition for a globally stable polymorphic equilibrium.

Let us assume a infinite number of players in a population, each player playing constantly strategy A, the altruist strategy. If the payoff values of this game are survival chances, and if this game is played an infinite number of times, a small group of N - players can successfully invade this population if and only if their chances to survive are strictly greater than the survival chances of the A - players. Equivalently, a small group of A - players can invade an infinite large population of N - players, if and only if their survival chances are strictly greater than the ones of the N - players. If we assume that the initial numbers of invading players is so small that a random mating of two invading players never occurs, we can state these conditions

more formally:

A can invade a pure N - player population if

$$(1 - E_2)u > (1 - E_1) \quad (7)$$

and N can invade a pure A - player population, if

$$(1 - E_2)(1 - u) > (1 - E_3) \quad (8)$$

We know from the assumptions of the model

$$(1 - E_2)u + (1 - E_2)(1 - u) > (1 - E_1)$$

and that ( altruism criterium ! )  $0 < u < 1/2$

so 
$$(1 - E_2)u > (1 - E_1)/2 \quad (9)$$

or 
$$u > \frac{1 - E_1}{2(1 - E_2)} \quad (10)$$

Likewise we know from the assumptions of the model and our condition (8)

$$(1 - E_2)u + (1 - E_2)(1 - u) > (1 - E_3) \quad (11)$$

and since  $1/2 < (1 - u) < 1$ , we are able to write

$$(1 - E_2)(1 - u) > (1 - E_3)/2$$

or 
$$u < \frac{E_3 + 1 - 2E_2}{2(1 - E_2)} \quad (12)$$

Both conditions together yield BOORMAN/LEVITT's condition for a globally stable polymorphism, which we can thus restate in the following way that a

globally stable polymorphism is possible if and only if both a pure N - population can be invaded by a small group of A - players, and a pure A - population by a small group of N - players. This inability of pure populations to resist invasion by players of the opposite strategy can be equivalently expressed as a impregnability of a certain mixed strategy - which is then equivalent to the concept of a stable polymorphism. This necessary and sufficient criterium must be equivalent to the criterium of the convexity of the differential extinction function, which we can from now justifiedly call fitness transfer function.

We may add that this double criterium does not preclude the possibility of periodic solutions .

### III

With this result we can now go on to the important findings made available through the COHEN/ESHEL (1976) and the MATESSI/JAYAKAR (1976) models. I will use as my point of departure the COHEN/ESHEL model, but will sometimes use MATESSI/JAYAKAR's notations. The purpose of these models is to show how even without genetic drift ( = a chance of fixation even for a recessive gene) and without the basic concept of group selection ( = differential extinction rates for whole demes ) genuine altruism can evolve. They still let this happen in small founder populations - for reasons which soon will become clearer.

Assumed is a infinite large haploid population of two types, the altruist A and the non-altruist E. Individuals mate randomly in the large population at the end of life cycle, but all other activities take place in small



groups of individuals with a finite number of members. Reproduction occurs at discrete and non overlapping generations, and so three stages of the life cycle can be distinguished:

- stage 1: small groups of individuals colonize empty habitats;
- stage 2: selection occurs in the form of differential viability/fecundity of individuals ( Mendelian selection) within the group.
- stage 3: individuals disperse and mate at random in the large population.

If we furthermore assume that the initial size of every colony is fixed and equals  $N$ , and that the individuals of the two different types A and E have equal chances of being included in each colony, we can describe the probability distribution for A and E types with the well known formula

$$a(m, N - m) = \binom{N}{m} p^m q^{N-m}$$

$p$  and  $q = 1 - p$  being the relative frequencies of the A and E type, respectively, in the entire population, and 'a' being the expected proportion of founding population containing exactly  $m$  A's and  $(N - m)$  E's.

Altruism, is defined:

- the presence of A's in a founder population increases the average rate of reproduction in the colony,
- the individual reproduction rate of an A individual is smaller than the one of an E individual

$h(p)$  and  $g(p)$  being the average numbers of A's and E's emerging from a single colony at the end of one period,  $p$  being the initial frequency of the A's.

After one period the relative frequency of A's in the entire population (denoted as  $p_1$ ) is :

$$p_1 = h(p)/(h(p) + g(p)) = f(p)$$

We know that  $h(0) = g(1) = 0$  and therefore  $f(0) = 0$  and  $f(1) = 1$  with  $p = 0$  and  $p = 1$  the two obvious equilibrium points of the system.

Since we regard discrete time periods, the stability criterium for  $p = 0$  is  $f'(0) < 1$ , and for  $p = 1$  it is  $f'(1) < 1$

$$\text{now } f'(p) = \frac{h'(p)[h(p) + g(p)] - h(p)[h'(p) + g'(p)]}{[h(p) + g(p)]^2} \quad (17)$$

which yields  $f'(0) = h'(0)/g(0)$

with the stability condition  $h'(0) < g(0)$

and  $f'(1) = -g'(1)/h(1)$

with the stability condition  $-g'(1) < h(1)$

Since (18)

$$h'(p) = \sum_{m=1}^N m \binom{N}{m} \phi_{a(m)} \left[ mp^{m-1} q^{N-m} + (N-m)p^m q^{N-m-1} \right]$$

and (19)

$$g'(p) = \lambda N \sum_{m=0}^{N-1} \binom{N}{N-m} \phi_{a(m)} \left[ mp^{m-1} q^{N-m} + (N-m)p^m q^{N-m-1} \right] - \lambda h'(p)$$

If we denote with  $\phi_a(m)$ , ( $m = 1, 2, \dots, N$ ) and  $\phi_b(m)$ , ( $m = 0, 1, \dots, N - 1$ ) the total value by which an A and an E, respectively, reproduces himself in a colony with  $m$  A's, we can state these two characteristics as

a)

$$m \phi_a(m) + (N - m) \phi_b(m) \geq (m - 1) \phi_a(m - 1) + (N - m + 1) \phi_b(m - 1)$$

or: 
$$\bar{\Phi}(m) \geq \bar{\Phi}(m - 1) \quad (13a, 13b)$$

(m) the average growth rate in a colony containing  $m$  A's, and the sharp inequality

$$N \phi_a(N) > N \phi_b(0)$$

clearly  $\phi_a(0)$  and  $\phi_b(N)$  have no meaning.

b) 
$$\phi_a(m) < \phi_b(m) \quad \text{for all } m \quad (14a)$$

If we assume a fixed ratio of reproduction rates between A's and E's, we can replace (2a) by:

$$\phi_b(m) = \lambda \phi_a(m), \text{ with } \lambda > 1 \quad (14b)$$

We can now describe the entire population by the following dynamic system

$$h(p) = \sum_{n=1}^N m \binom{N}{m} \phi_a(m) p^m q^{N-m} \quad (15)$$

$$g(p) = \sum_{m=0}^{N-1} (N-m) \binom{N}{N-m} \phi_b(m) p^m q^{N-m}$$

$$g(p) = \lambda N \sum_{n=0}^N \binom{N}{N-m} p^m q^{N-m} \phi_a(m) - \lambda h(p) \quad (16)$$

and if we furthermore assume  $N$  to be sufficiently large\* so that the existence of one  $A$  is compatible with setting  $p = 0$ , or likewise that the existence of one  $E$  is compatible with setting  $p = 1$  then the stability criterium for  $p = 0$  becomes

$$h'(0) = \varphi_a(1)N$$

$$g(0) = \varphi_b(0)N$$

or

$$\varphi_a(1) < \varphi_b(0)$$

Similarly for  $p = 1$

$$-g'(1) = \lambda \varphi_a(N-1)N = \varphi_b(N-1)N$$

$$h(1) = \varphi_a(N)N$$

or

$$\varphi_b(N-1) < \varphi_a(N)$$

From these two stability criteria we can argue, that a necessary condition for a globally stable polymorphism is that both edge-equilibria are unstable, i.e.

$$\varphi_a(1) > \varphi_b(0) \quad (20a)$$

and

$$\varphi_a(N) < \varphi_b(N-1) \quad (20b)$$

If we recall the assumption  $\varphi_b(m) = \varphi_a(m)$ , then we can rewrite these two conditions as

$$\frac{\varphi_b(1)}{\varphi_b(0)} > \lambda > \frac{\varphi_a(N)}{\varphi_a(N-1)} \quad (20c)$$

we may add that  $\lambda$  is not necessarily a constant which, for purposes of simplicity, we have treated it thus far. We can then interpret this important condition:

for appropriate values of  $\lambda$  a stable polymorphism is possible if the relative

---

\*A necessary condition not made explicit by COHEN/ESHEL nor MATESSI/JAYAKAR.

benefit a founder population of non altruists receives from a single altruist exceeds the relative damage a founder population of altruists suffers from a single non altruist.

A second look into the inequality  $\phi_a(N) > \phi_b(N-1)$ , the precondition for a stable monomorphic altruist equilibrium, clearly demonstrates how little the likelihood for such a equilibrium is.

We can express this inequality as

$$\frac{\phi_a(N)}{\phi_a(N-1)} > \lambda \text{ as well as } \frac{\phi_b(N)}{\phi_b(N-1)} > \lambda \quad (21a, 21b)$$

and 
$$\frac{\phi_a(N)}{\phi_b(N-1)} > 1 \quad (21c)$$

In real life situations  $\phi_a$  can beyond a certain N only be a decreasing function, so if we let N increase

$$\lim_{N \rightarrow \infty} \frac{\phi_a(N)}{\phi_a(N-1)} \leq 1 \quad (22)$$

But by definition  $\lambda > 1$ , so N cannot grow without boundaries. But condition (21c) can only be met if and only if  $\phi_b$  is not only a monotone decreasing function, but decreases even faster than  $\phi_a$ , which is hardly imaginable in any real life situation.

It is easy to see that in order to make a unique and globally stable equilibrium possible,  $\phi_a(m)$ , ( $m = 1, 2, \dots, N$ ) must be a monotone decreasing function, and  $\phi_b(m)$ , ( $m = 0, 1, \dots, N-1$ ) must be a monotone increasing function.

If we recall the first formal condition of altruism (13a, 13b)

$$m\phi_a(m) + (N-m)\phi_b(m) \geq (m-1)\phi_a(m-1) + (N-m+1)\phi_b(m-1)$$

$$\Phi(m) \geq \Phi(m-1)$$

we can easily see that given these two properties of the  $\phi_a(m)$  and the  $\phi_b(m)$  function, clearly  $\Phi(m)$  must be convex. This function  $\Phi(m)$  can be interpreted as the average fitness transfer function, i.e. the fitness received by the members ( altruists and non altruists alike ) of a founder populations from the presence of  $m$  altruists in this population.

If  $\Phi(m)$  is linear or concave, at least one edge ( $p = 0$  or  $p = 1$ ) is stable, but no stable polymorphism is possible.

The model here, based on COHEN/ESHEL (1976) and MATESSI/JAYAKAR (1976), still shares with the classical group selection models the idea that altruism can develop only in small founder populations, since only in very small groups can the altruists increase in numbers so rapidly that they can compensate with increased fitness transfer for their individual handicap vis-a-vis the non altruists. In another version of the group selection approach, which assumes a population permanently divided in small founder colonies and has therefore to allow some migration between colonies, a sufficiently low migration rate is also essential for the evolution of altruism (GILPIN 1975)

This assumption of at least temporary dispersal into small founder demes can easily be replaced by the notion of some phenotypic assortment of encounters, which we can define as any deviation from encounter randomness which is not based on kinship. Altruists might tend to congregate. This idea can be refined even more by the idea of an active search of altruists for companions or the like. Formally one can grasp this in the concept of some internal viscosity of large continuous populations. Theoretical and empirical studies on the evolution

of cooperativeness have demonstrated the importance of this internal viscosity concept (AXELROD/HAMILTON 1981, ESHEL/CAVALLI-SFORZA 1982, BOORMANN/LEVITT 1980, ch. 2 - 5). I expect the same for our subject, the evolution of altruism. It is clear, however, that replacing the notion of founder populations with the idea of the internal viscosity of the large population eliminates the very possibility of a group selection approach, for which the idea of differential group extinction is essential.

#### IV

In order to interpret these findings of population biology I wish first to mention two things:

1) The notion of a stable polymorphism of different players pursuing different pure strategies can without any problems be substituted for by the concept of an equilibrium of a mixed strategy applied uniformly by all the players in the game ( under the requirement that every player faces the same payoff matrix - as it was the case in the models presented above).

By this we are replacing the impregnability of a polymorphism of actors by the impregnability of a mixed strategy - this is the essence of MAYNARD-SMITH's (1973, 1974, 1976) well known concept of a Evolutionary Stable Strategy (ESS). An ESS is defined as a strategy which, if adopted uniformly by a population, cannot be invaded by an initially rare mutant adopting another strategy.

The formal requirements for a strategy to be an ESS is, if  $E_j(I)$  denotes the benefit for I yielded by an "encounter" (which stands for all sorts of

interactions) with J, the I is ESS, if for all alternative strategies J

$$E_I(I) > E_I(J) \quad (23)$$

or  $E_I(I) = E_I(J) \quad \text{and} \quad E_J(I) > E_J(J) \quad (24)$

Back to our altruism evolution model, where a stable polymorphism is characterized such that there is a  $m_s$  for which the fitness transfer function  $\bar{\Phi}(m_s)$  has its maximum, so that

$$\bar{\Phi}(m_s - 1) < \bar{\Phi}(m_s) > \bar{\Phi}(m_s + 1) \quad (25)$$

If  $\bar{\Phi}(m)$  is convex between N and 0, this maximum is unique.

In terms of MAYNARD-SMITH's concept this  $\bar{\Phi}(m_s)$  can be translated as an (mixed) ESS, which requires to play with relative frequency  $p = m_s / N$  the altruist strategy and with relative frequency  $q = (N - m_s) / N$  the non altruist strategy.

2) Secondly, we can easily separate these ideas of a stable polymorphism and a mixed ESS from the original basis in genetics.

One has only to set up a learning model which is endowed with both the altruist and the non altruist strategy, but can optimize the relative frequency with which the one or the other can be applied.

The conclusions which we can draw from all this are the following:

We were not only able to specify the conditions under which the evolution of altruism qua group selection can occur, we were able to redefine all collective parameters, referring to the ecological, spatial and other structural properties of the population as a whole, as parameters of individual behavior strategies. For reasons already mentioned, we focused on



the conditions for survival and development of a stable polymorphism of altruism/non altruism. We could then replace the convexity of differential group extinction function with the convexity of the individual benefit transfer function, replacing the requirement regarding the disperse structure of the relevant populations with ideas of non-random assortment, low migration, and active search by altruists for other altruists (briefly: the concept of internal viscosity of a population), and finally replacing the idea of the impregnability of a stable polymorphism in a population with the idea of a mixed ESS.

And these redefinitions are not just a conceptual convenience, since group selection can not account for a development of altruism in large continuous populations. There is broad evidence from empirical studies which limits the relevance of group selection as a driving force of evolution even more, beyond this crucial point. (MAYNARD-SMITH 1976, WADE 1978, BOORMAN/LEVITT 1980) The models presented here do not only convincingly suggest to conceive the development of altruism in terms of evolution of individual behavior rationality, they give us an important hint how to do this.

I mean the inequality:

$$\frac{\varphi_b(1)}{\varphi_b(0)} > \lambda > \frac{\varphi_a(N)}{\varphi_a(N-1)} \quad (20c)$$

which connects the two elementary fitness transfer functions

$\varphi_a(m)$  and  $\varphi_b(m)$  with the parameter  $\lambda$  which describe the individual disadvantage of an altruist vis-a-vis a non altruist.

We have hitherto regarded  $\lambda$  as a constant, but this of course does not need to be so; in fact, we have to consider it as the other important function in

any model of evolutionary altruism, besides the benefit-transfer function. Optimizing the average fitness or - for those social scientists who frown upon this very word - the average welfare in a society which is created by the existence of altruists, means maximizing the benefit transfer function and minimizing the  $\lambda$  function, as we now can call it.

The understanding of such an optimization process will provide us with the key to understand how human societies are, and in fact were able to solve the free rider problem. And the essence of this approach is that we have transformed the free rider problem from the level of institutions to the level of evolutionary stable strategies of individuals.

v

OLSON himself has stated his theorem in the cautious form that in groups where a single individual is ready to pay for the public good himself alone if necessary, the group will normally provide it, and that in a group in which a single person's contribution will not make a noticeable difference in the amount of the public good provided to the group, no supply of the public good will occur at all. For the size range between his theorem does not allow predictions. Supply can occur or not. OLSON points to the theory of games in order to clarify this problem (OLSON 1971, 43), and indeed this is also the conceptual framework, to which our previous findings lead us.

The free rider situation can be represented by a N - person prisoner's dilemma, preferably in compound form (HAMBURGER 1973), i.e. it is assumed that each player plays the same strategy simultaneously against all other N - 1 players, so that for each player the N - person game can be analytically

partitioned into  $(N - 1) 2 -$  person prisoner's dilemmas like:

		2. player	
		cooperate	defect
1. player	cooperate	A <sub>1</sub> A <sub>2</sub>	B <sub>1</sub> C <sub>2</sub>
	defect	C <sub>1</sub> B <sub>2</sub>	D <sub>1</sub> D <sub>2</sub>

with  $C_i > A_i > D_i > B_i$  ( $i = 1, 2$ );  $A_1 = A_2$ ,  $B_1 = B_2$ ,

$C_1 = C_2$ ,  $D_1 = D_2$ ; and  $A > (B + C)/2$ .

If we ask in this context how the  $\Phi$  function under the restraint of  $\lambda$  can be maximized, we know that the  $\Phi$  function cannot be a linear payoff function, if we want to see a stable polymorphism possible. One of the consequences of that is that the  $\Phi$  function can not be dependent on a single move, but on a strategy over time.

Strategy shall mean a decision rule which is a function from the previous history ( the whole history or only the last 1,2,3..... moves) of the game into the next move, i.e. derives from the relevant chain of already made moves and their payoffs the next move or moves.

It is well known that the time of the game, expressed in the number of moves, in the context of the prisoner's dilemma, must not be predetermined and known

to the players in advance if cooperation between the players shall remain a rational option at all (LUCE/RAIFFA 1957, 97 - 102). The only way for, say, the first player to enforce cooperation (moves: cooperate - cooperate) and to prevent (cooperate - defect) is the threat of retaliation (defect - defect) or even (defect - cooperate). If the number of moves is fixed, say N, and known to the players in advance, then at the N-th move each player must defect, since for a possible defection of the other player he has no chance of retaliation any more. So the N-th move must be a (defect - defect) move. But then there will be no chance of a retaliation for the (N - 1)-th move, so the rational choice for this move must also be (defect - defect) and so forth all the way down to the first move. Therefore the number of moves must be undetermined. But how then to compute the payoff of a strategy which consists of an indefinite, maybe even an infinite number of moves.

The familiar solution to this question is introducing a discount, preferably a constant one, so that for example the present value of a strategy of permanent mutual defections for each player would be:

$$\begin{aligned} V &= D + wD + w^2D + w^3D + \dots \quad (26) \\ &= D/(1 - w) \end{aligned}$$

In a series of excellent articles, from which I will draw also at some further points, AXELROD (AXELROD 1980a,1980b,1981; AXELROD/HAMILTON 1981) has described the two possible interpretation of this discount parameter 'w'. The standard interpretation is that future consumption is not as valuable as present consumption.

The other interpretation is that future moves will eventually not occur, because the players won't meet for another encounter any more. It is very easy to see in this interpretation the dicount parameter as an parameter for the internal viscosity of the population, in which this game takes place. We can

then easily let  $N$ , the number of players be infinite. AXELROD (1981) can prove several theorems, three of which I will use here for my further argumentation:

1) If the viscosity parameter is sufficiently high, there is no best strategy independently of the strategy used by the other player. The proof for this is important enough to be presented here (slightly amplified).

Proof: Suppose  $X$  is a best strategy independently of the other player's strategy. That means for every other strategies  $X'$  and  $Y$  the value of strategy  $X$  played against  $Y$  must be higher than the value of  $X'$  played against  $Y$  :  $V(X/Y) > V(X'/Y)$ . Let us define a strategy to be nice, if it never defects first. Let us now consider first the case that  $A$  is nice. Let  $X' = \text{ALL D}$ , a strategy of permanent defection, and  $Y$  be  $\text{ALL C}$ , a strategy always to cooperate. Then  $V(X/Y) = A/(1 - w)$  which is smaller than  $V(X'/Y) = C/(1 - w)$ . So there is a strategy  $X'$  which does better against some strategy  $Y$  than any nice strategy.

Suppose  $X$  is not a nice strategy. Then let  $X' = \text{ALL C}$  and let  $Y$  be a strategy which always cooperate as long as the other player cooperates, but defects forever as soon as the other player defects the first time. If  $X$  is not a nice strategy, it will at some point defect as first, say after  $n$  moves. Then  $V(X/Y) \leq [A + w^{n-1}A + w^n(C + wC + wD)]/(1 - w)$ .  $V(X/Y)$  is maximal if  $n = \infty$ , then  $V = R / (1 - w)$ .  $V(X/Y)$  is minimal when  $n = 1$ , then  $V = C + wD/(1-w)$ . But  $V(X'/Y) = A/(1 - W)$ . In this case  $V(X/Y) < V(X'/Y)$  whenever  $w > (C - A)/(C - D)$ . So there is for every not nice strategy  $X$  a strategy  $X'$  and a strategy  $Y$  such that  $V(X/Y) < V(X'/Y)$ , provided  $w$  is sufficiently large. So there is no strategy which is strictly better than any other strategy, regardless of the strategy used by the other player. q.e.d.

AXELROD favours therefore an evolutionary approach for finding optimal strategies in the sense of MAYNARD-SMITH's ESS (actually AXELROD narrows the original definition of ESS to the inequality  $V_I(I) > V_I(J)$ , and calls it a collectively stable strategy). Such a collectively stable strategy is in NASH equilibrium with itself. An optimal strategy is then a strategy which cannot be invaded.

Pursuing this evolutionary approach AXELROD can furthermore prove:

- 2) Any strategy which is the first one to cooperate - either at the first move or later - can be invaded only if  $w < (C - A)/(C - D)$ .
- 3) Small clusters of players using nice strategies can invade ALL D strategy.

In order to make these findings useful for our own problem we have to take a second look into the viscosity parameter 'w', which we have treated as a constant so far. Now in an infinite population there will be people which I will hopefully meet time and again, but there will be a large number of people, for whom the chances of a second encounter with me are arbitrarily small. So 'w' is in fact not some constant nor a simple variable, but stands for a probability distribution which ranges from some maximum value close to one to almost zero.

Very tentatively I will now state four conjectures:

- 1) Nice strategies, which are altruist strategies, can in small clusters invade non altruist strategies. But in an infinite population they are not able to crowd the non altruist out completely, because there is always a segment of the set of players, for which the viscosity parameter is small enough to allow a non altruist to invade the altruist strategy. So in a infinite population, in which clustering is permitted, any ESS must be a mixed one.

- 2) If we define an average welfare function for such a population, in which the ESS is mixed, this function must be convex.
- 3) There may be no best altruist strategy, but perhaps it is possible to find properties of a class of best altruist strategies which do equally better against all other altruist strategies, but are not able to dominate each other.
- 4) The evolution of altruism depends on the development of strategies which can invade non altruist strategies in minimal clusters, can prevail against non altruist strategies at minimal viscosity, and will for a given encounter probability distribution minimize the proportion of the average welfare which falls to the non altruists.

## VI

Three lines of experimental research which are important for our subject, can be distinguished.

- 1) The relatively small number of studies which directly try to verify the SAMUELSON/OLSON theorem, testing according hypotheses on the influence of group size, distribution of assets, effect of provision points and the like, on the actual contribution of the group to a public good (BOHM 1971; MARWELL/AMES 1979, 1980; SWEENEY 1973). While providing some relevant detail findings, the relevance of these studies is limited by their flawed theoretical framework: they expect manifest free rider behaviour, and take its absence as an empirical disconfirmation of the theorem.
- 2) The "Social Loafing" research (HARKINS/PETTY 1982; KERR/BRUUN 1981, 1983; LATANE/WILLIAMS/HARKINS 1979; WILLIAMS/HARKINS/LATANE 1981) These studies

focus on the effect of perceived dispensability and identifiability of the individual contribution on efforts and achievements in group problem solving tasks. While some of the experimental findings are extremely important - for example the exceptional role of identifiability as a prerequisite for effective sanctions against free riding - the conceptual framework does not pay any attention to rational choice nor to evolutionary approaches, but instead interpretes these findings in the light of psychological theories of small range.

- 3) The vast body of experimental Prisoner's Dilemma studies (recent studies with a good bibliography are CALDWELL 1976; KOOMORITA/SWEENEY/KRAWITZ 1980). Closely related is the research on "Social Dilemmas"(DAWES 1980; KOOMORITA/LAPWORTH 1982) The focus of the experimental Prisoner's Dilemma studies is on what kind of strategies people actually apply. A special version is the simulation of tournaments, competing strategies are confronted with each other in order to assess their relative strength (AXELROD 1980a, 1980b).

We can assume that , through life experiences, through internalized cultural patterns, or even through some anthropological endowments, people have some implicit knowledge about such strategies and how to use them in accordance to the environment and have certain subjective expectations concerning the strategies of other players. These expectations can be modelled as subjective probability distributions over the possible strategies employed by other players. In the experimental study we are conducting right now we investigate two aspects of these implicit expectations.



Secondly we are investigating how the actual content and the reliability of information about other players shape the single actor's subjective expectations. The content simply refers to the actual contributions each player is likely to make for the acquisition of the public good. The higher the contributions of other players, the closer to the optimum point the single actor's contributions should be. Reliability of information can be defined as the relative level of information within a group which is available to single member, and can be measured as the ratio between the information, which is actual available to a member and the information about the group, which could at most be made available to him. We consider only information on the willingness of other members to contribute to the public good.

So our second hypothesis is that the lower the relative level of information is for a decision maker, the less confident he will be that none of the other players will take a free ride, and therefore will have to make a contribution further away from the optimum point: either to accept a disproportionately high share of the total costs, or to contribute nothing. On the other hand, the higher the relative level of information for one actor, the closer will be his contribution to the optimum point.

In order to test this hypotheses we have designed a simple computer simulation program. Our subjects are communicating with four virtual partners about the sharing of the costs of a public good, so that we are simulating the communication in a five member group. We have devised two different structures

The basic idea is the following:

If we assume equal interests in the supply of some public good and equal initial assets among the players, then the optimum point for the supply of this good is the same for each player, and equals that supply where the marginal utility of this public good equals the marginal price per player. The more likely it is that one or even more of the players might take a free ride, the more must a rational player's contribution for the purchase of the public good deviate from this optimal point - either the contribution must be higher in order to compensate for the loss of the free rider's contribution, or it must drop to the level of a non contribution as well, if the marginal utility of the public good never exceeds the marginal price.

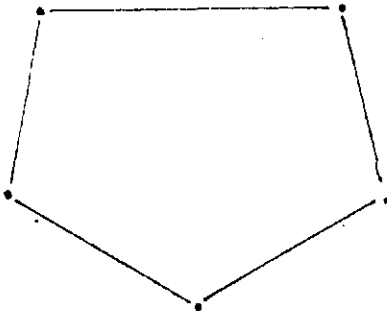
Two important related aspects of the development of the subjective expectations are the player's own risk aversiveness and the player's assumptions about the other players' aversiveness. We want to make use of the well known fact, recently thoroughly studied by KAHNEMANN/TVERSKY (1978, 1979, 1981), that in terms of the conventional expected utility theory, people tend to be risk averse if the issue is how to realize a possible gain, and tend to be risk taking if the issue is how to avoid a possible loss.

Clearly, if no player can be forced to refrain from taking a free ride, going closer to the optimality point in one's own contribution is risky.

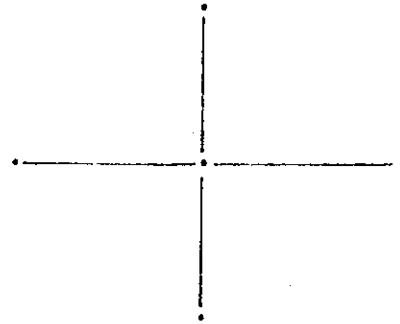
So our first hypothesis is that in how-to-avoid-a-possible-loss situation people will tend to make contributions closer to the optimum point than in a how-to-realize-a-possible-gain situation. This hypothesis refers to two mechanisms: the risk taking behaviour of the actor himself, and the unconscious or conscious expectations that the other players may behave in the same way.

for this communication,

a ring structure, and



a star structure.



These structures allow three different positions: one of the equivalent positions in the ring structure, one of the equivalent peripheral positions in the star structure, and finally the central position in the star. In our simulation each player can communicate only with his immediate neighbours, so in a peripheral position in the star you have just one communication partner, in the ring two, and in the peripheral in the star position four, with relative information levels of  $1/4$ ,  $1/2$  and  $1$ , respectively.

A third focus in our experiment is on whether experiencing other partners' defections in the past has an impact upon later decisions. We hypothesize that information about former defections in the group will make players more likely to make contributions further away from the optimum point.

Further details of the experimental design will be presented at the meeting, together with first results, which will hopefully be available by then.

Bibliography

- AXELROD, R. 1980a  
Effective Choice In The Prisoner's Dilemma.  
J. of Conflict Resolution 24, 3 - 25 (1980)
- AXELROD, R. 1980b  
More Effective Choices In The Prisoner's Dilemma.  
J. of Conflict Resolution 24, 379 - 403 (1980)
- AXELROD, R. 1981  
The Emergence Of Cooperation Among Egoists.  
American Political Science Review 75, 306 - 318 (1981)
- AXELROD, R.; W.D. HAMILTON 1981  
The Evolution Of Cooperation.  
Science, 211, 1390 - 1396 (1981)
- BOHM, P. 1972  
Estimating Demand For Public Goods: An Experiment.  
European Economic Review 3, 111 - 130 (1972)
- BOORMAN, P.; S. LEVITT 1980  
The Genetics Of Altruism.  
Academic Press 1980
- CALDWELL, M.D. 1976  
Communication And Sex Effects In A Five-Person Prisoner's Dilemma Game.  
J. of Personality and Social Psychology 35, 273 - 280 (1976)
- CLARKE, E.H. 1971  
Multipart Pricing Of Public Goods.  
Public Choice 11, 17 - 33 (1971)
- COHEN, D.; I. ESHEL 1976  
On The Founder Effect And The Evolution Of Altruistic Traits.  
Theoretical Population Biology 10, 276 - 302 (1976)
- DAWES, R.M. 1980  
Social Dilemmas.  
Annual Review Of Psychology 31, 169 - 193 (1980)
- ESHEL, I. 1972  
On The Neighbor Effect And The Evolution Of Altruist Traits.  
Theoretical Population Biology 3, 258 - 277 (1972)
- ESHEL, I. 1982  
Evolutionary Stable Strategies And Viability Selection In Mendelian  
Populations.  
Theoretical Population Biology 22, 204 - 217 (1982)

- LUCE, R.D.; H. RAIFFA 1957  
Games And Decisions.  
John Wiley 1957
- MARWELL, G.; R.E. AMES 1979  
Experiments On The Provision Of Public Goods I: Resources, Interest,  
Group Size, And The Free-Rider Problem.  
American Journal of Sociology 84, 1335 - 1360 (1979)
- MARWELL, G.; R.E. AMES 1980  
Experiments On The Provision Of Public Goods II: Provision Points,  
Stakes, Experience, And The Free-Rider Problem.  
American Journal of Sociology 85, 926 - 937 (1980)
- MATESSI, C.; S.D. JAYAKAR 1976  
Conditions For The Evolution Of Altruism Under Darwinian Selection.  
Theoretical Population Biology 9, 360 - 387 (1976)
- MAYNARD-SMITH, J. 1974  
The Theory Of Games And The Evolution Of Animal Conflicts.  
J. of Theoretical Biology 47, 209 - 221 (1974)
- MAYNARD-SMITH, J. 1976  
Group Selection.  
Quart. Review of Biology 51, 277 - 283 (1976)
- MAYNARD-SMITH, J.; G.R. PRICE 1973  
The Logic Of Animal Conflict.  
Nature 246, 15 - 18 (1973)
- OLSON, M. 1971  
The Logic Of Collective Action.  
Harvard University Press 1971
- SWEENEY, J.W. 1973  
An Experimental Investigation Of The Free-Rider Problem.  
Social Science Research 2, 277 - 292 (1973)
- TIDEMAN, T.N.; G. TULLOCK 1976  
A New And Superior Process For Making Social Choices.  
J. of Political Economy 84, 1145 - 1159 (1976)
- TVERSKY, A.; D. KAHNEMAN 1981  
The Framing Of Decisions And The Psychology Of Choice.  
Science 211, 453 - 458 (1981)
- WADE, M.J. 1978  
A Critical Review Of The Models Of Group Selection.  
Quart. Review of Biology 53, 101 - 114 (1978)
- WILSON, D.S. 1975  
A Theory Of Group Selection.  
Proc. Nat. Acad. Sci. USA 72, 143 - 146 (1975)

- ESHEL, I.; L.L. CAVALLI-SFORZA 1982  
Assortment Of Encounters And Evolution Of Cooperativeness.  
Proc. Natl. Acad. Sci. USA 79, 1331 - 1335 (1982)
- GILPIN, M.E. 1975  
Group Selection In Predator-Prey Communities.  
Princeton University Press 1975
- GROVES, Th.; J. LEDYARD 1977  
Optimal Allocation Of Public Goods: A Solution To The "Free Rider"  
Problem.  
Econometrica 45, 783 - 809 (1977)
- HALDANE, J.B.S. 1932  
The Causes Of Evolution.  
Longmans/Green 1932
- HAMBURGER, H. 1973  
N - Person Prisoner's Dilemma.  
J. of Mathematical Sociology 3, 27 - 48 (1973)
- HAMILTON, W.D. 1964a, 1964b  
The Genetical Evolution Of Social Behaviour. I + II  
J. of Theoretical Biology 7, 1 - 52 (1964)
- HARKINS, S.G.; R.E. PETTY 1982  
Effect Of Task Difficulty And Task Uniqueness On Social Loafing.  
J. of Personality and Social Psychology 43, 1214 - 1229 (1982)
- KAHNEMAN, D.; A. TVERSKY 1979  
Prospect Theory: An Analysis Of Decisions Under Risk.  
Econometrica 47, 263 - 291 (1979)
- Kahnemann, D.; A. TVERSKY 1982  
The Psychology Of Preferences  
Scientific American Jan 1982, 160 - 173
- KERR, N.L.; S.E. BRUUN 1981  
Ringelmann Revisited: Alternative Explanations For The Social Loafing  
Effect.  
Personality and Social Psychology Bulletin 7, 224 - 231 (1981)
- KERR, N.; S.E. BRUUN 1983  
Dispensability Of Member Effort And Group Motivation Losses: Free-Rider  
Effects.  
J. of Personality and Social Psychology 44, 78 - 94 (1983)
- KOMORITA, S.S.; C.W. LAPWORTH 1982  
Alternative Choices In Social Dilemmas.  
J. of Conflict Resolution 26, 692 - 708 (1982)
- KOMORITA, S.S.; J. SWEENEY; D.A. KRAVITZ 1980  
Cooperative Choice In The N - Person Dilemma Situation.  
J. of Personality and Social Psychology 38, 504 - 516 (1980)

WILSON, D.S. 1979

Natural Selection Of Populations And Communities.  
Benjamin/Cummings 1979

WYNNE-EDWARDS, V.C. 1963

Intergroup Selection In The Evolution Of Social Systems.  
Nature 200, 623 - 626 (1963)