STRUCTURED DEMES AND THE EVOLUTION OF GROUP-ADVANTAGEOUS TRAITS

DAVID SLOAN WILSON*

Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138

I. THE CONCEPT OF STRUCTURED DEMES

Earlier (D. S. Wilson 1975) I presented a theory for the selection of "altruistic" traits, denied by traditional models of individual selection. Its mechanism and many of its predictions also differ from traditional models of group selection. Charnov and Krebs (1975) and Hamilton (1975) simultaneously built similar models.

Here I examine the ingredients necessary for this theory to operate and the extent to which they may be expected to occur in nature. Sections II and III explore some of the implications for ecological and evolutionary models.

Demes and Trait Groups

Evolution's most easily conceived population unit is the deme, which can be idealized as "a group of individuals so situated that any two of them have equal probability of mating with each other and producing offspring" (Mayr 1963, p. 136). The same concept applies to haploid and two-species associations. Fundamentally we are interested in the set of organisms with equal probability of interacting with each other.

In determining deme size, care must be taken to inspect the entire life cycle of the organism. Individuals usually interact with their closest neighbors, but if at some other point in the life cycle widespread movements occur (such as a dispersal stage) the deme is much larger than such neighbor interactions would suggest. Thus an equivalent definition for the deme would be a group of individuals that readily intermix during some point in their life cycle, giving any two an equal probability of becoming neighbors.

This, however, says nothing about how individuals interact during their life cycle. For this an ecological analogue to the deme concept is needed—a group of individuals that readily interact with each other in competition, aggression, mating, defense against predation, or any other process of ecological interest. Examples come easily: a vessel-inhabiting mosquito larva interacts only with others within the same vessel. A territorial animal deals only with its neighbors.

* Present address: Department of Zoology, University of Washington, Seattle, Washington 98195.

A plant competes for nutrients only within the space of its root system and for light only within the area of its crown.

These "ecological demes" are functions not only of the organisms and their habitats but also of the particular traits being manifested. A caterpillar competes with others only on the same plant, but as a butterfly it mates and competes with a much larger group. Similarly, a school of fish is a closed "ecological deme" with respect to predation avoidance, but several neighboring schools may compete for food. Accordingly, I have termed ecological demes "trait groups"—groups of individuals within which interactions readily occur with respect to the trait under consideration (D. S. Wilson 1975).

In general, two forms of trait groups may be recognized: (1) discrete, in which the organisms are actually partitioned into isolated habitat units such as tree holes, dung pads, fungus fruiting bodies, or parasitized host for parasites; and (2) continuous, in which individuals are distributed relatively evenly over the deme but each interacts only with its immediate neighbors. Each individual thus forms the center of its own trait group, and trait groups overlap with each other. Examples are territorial animals and most plants. Discrete and continuous trait groups are the analogues to the isolated deme and the "neighborhood" of population genetics.

What is the relation between the trait group and the deme? Trait groups are usually much smaller, and for a very general reason. Most organisms tend to concentrate their movements in a brief dispersal state—the seeds and pollen of plants, the postteneral migrating stage of insects, the larvae of benthic marine fauna, and the adolescents of many vertebrates. At all other periods in the life cycle the individuals are relatively sedentary, their movements trivial in scale compared to their dispersal period. The short burst of intensive movement sets the size of the deme. It follows that any traits manifested during the nondispersal stage will be within trait groups smaller than the deme.

The entire theory presented here depends on this fact. In essence it means that many demes are not only a population of individuals but also a population of groups. The groups are isolated with respect to the manifestation of traits, but the individuals composing the groups periodically coalesce and re-sort due to the dispersal process. Such demes are not internally homogenous. They have a structure and may be termed "structured demes."

Much has been made recently of the smallness of demes (Ehrlich and Raven 1969; Levin and Kerster 1974). Bradshaw (1972) suggests for plants that "effective population size . . . is to be measured in meters and not kilometers." But this is still much larger than most trait groups. A leaf miner's trait group equals a leaf.

Consider, for example, the perennial herb *Phlox pilosa*. Levin and Kerster (1968) found sharply leptokurtic distributions for both seed and pollen dispersal. Pollen was transferred mainly to neighboring plants, and the average dispersal of seeds was only 1.1 m. This is slight but not atypical of many plant populations (Levin and Kerster 1974).

Levin and Kerster calculated the neighborhood size for *Phlox* to be 94–983

individuals covering an area of 11.6–88.4 m². Yet an individual plant would never be expected to actively compete for light or nutrients within such a large space. Even though its neighborhood is small, its trait group is still smaller.

Deme structure is especially pronounced among the insects (Johnson 1969). Shortly after mating, many insects embark upon an extended flight. Hardly any traits are manifested during dispersal—all "appetites" such as feeding and oviposition are suppressed (this is actually Johnson's definition of a dispersal flight). The distances traveled are orders of magnitude greater than larval movements or subsequent adult movements. Almost all trait groups are far smaller than the deme.

The concept of structured demes is thus quite general—perhaps general enough to be called a "rule," although exceptions do of course exist. Many animals, such as sea lions, congregate in dense colonies to mate and breed. The trait group for all such activities approaches the deme. Some organisms do not possess dispersal stages. Small isolated habitats may prevent dispersal. For example, fish in very small ponds are restricted in their movements. If each individual feeds throughout the entire pond, then this trait group will equal the deme. Finally, laboratory "bottle experiments" might provide the most common occurrence of unstructured demes. As exceptions, these cases may be good testing grounds for the theory presented here.

The concept of structured demes is certainly not new. It is similar to the subdivided populations of multiple-niche models (e.g., Christiensen and Feldman 1975) and also shares elements of Hamilton's "population viscosity" (Hamilton 1971). All these models deal with individuals that are somehow restricted to areas smaller than the deme. Multiple-niche models treat the situation where the genotypes differ in their niche requirements. This is a different subject and will not be considered further here. The relationship between kin selection and structured demes is treated in detail below.

Implications for ecological and evolutionary models.—The standard model of selection considers two types (alleles, genotypes, species) endowed with the desired properties present in the population at a given density and at frequencies a and $(1 - a)$. The model is then run to see which type is selected for or if both are maintained at equilibrium.

The unstated assumption of these models is that the demes are unstructured. Structured demes, whose members are isolated into trait groups, cannot be characterized by any single density or frequency, since these will vary among trait groups. The overall density and frequency of types in the deme (= average density and frequency per trait group) are not relevant parameters as far as the organisms are concerned. Lloyd (1967) showed that, given a variation in density, the subjectively experienced ("mean") crowding is invariably higher than the average density of the population. As will be shown here, the same holds true for frequency, and it is this fact that allows the evolution of altruistic traits.

Average and subjective frequencies.—Consider a deme subdivided into four discrete trait groups, as shown in figure 1. Twenty animals comprise the deme,

I A	2 A	3 A	4 A
4 B	3 B	2 B	I B

FIG. 1.—Four discrete trait groups that vary in the frequency of types

five in each trait group. Two types (A and B) are present in varying proportions among the trait groups. We are interested in the following parameters:

a, b = the average frequencies of the two types in the deme ($b = 1 - a$);

a_A, b_A = the frequency of the two types, experienced by the average A type (average subjective frequency);

a_B, b_B = the frequency of the two types, experienced by the average B type.

They are all easily calculated as follows:

$$a = b = \frac{.2 + .4 + .6 + .8}{4} = .5;$$

$$a_A = b_B = \frac{1(.2) + 2(.4) + 3(.6) + 4(.8)}{10} = .6;$$

$$b_A = a_B = \frac{1(.8) + 2(.6) + 3(.4) + 4(.2)}{10} = .4.$$

The "subjective" frequencies are simply the weighted averages over all the trait groups. In other words, only one A type experiences the low frequency of .2 A types (itself), but four A types experience the high frequency of .8 A types, and so on. The concept is exactly parallel to Lloyd's (1967) treatment of variations in density.

In general, assume T discrete trait groups of equal density (N), each with proportions a_i and b_i of the A and B types. (As before stated, in nature N will also vary among trait groups, but for modeling purposes it is best to begin with it constant. I have examined density variation in computer simulations for most of the models presented here [D. S. Wilson, in preparation] and find no qualitative change in the conclusions. Variations in density are also treated by Lloyd [1967] and S. A. Levin [1974], among others.) The subjective frequencies are then

$$a_A = \frac{\sum a_i^2}{\sum a_i} = a + \sigma^2/a, \qquad b_A = \frac{\sum a_i b_i}{\sum a_i} = b - \sigma^2/a$$

$$a_B = \frac{\sum a_i b_i}{\sum b_i} = a - \sigma^2/b, \qquad b_B = \frac{\sum b_i^2}{\sum b_i} = b + \sigma^2/b,$$

(1)

where σ^2 equals the variance in frequency among trait groups. This follows directly from Lloyd (1967) when the density in each trait group is held constant.

The subjective frequencies equal the average frequencies only when there is

zero variance among the trait groups, that is, complete uniformity. A nonzero variance—even the variance arising from a binomial distribution—acts to increase the subjective frequency of one's own type and depress that of the other.

Continuous trait groups are difficult to treat analytically but easy to simulate on a computer. Accordingly, a 105×105 matrix was set up and "filled" with A and B types. The types were entered at several overall frequencies ($a = .05$, .15, .4, .6, .85, .95) in distributions both patchy and random with respect to each other. First the matrix was divided into discrete trait groups and the subjective frequencies calculated. Trait groups were then constructed around every individual in the matrix to simulate the continuous case and again the subjective frequencies tabulated. Trait groups could not be constructed around edge individuals, which were omitted. Three trait-group sizes were used: 3×3, 5×5, and 7×7. In all cases the discrete and continuous trait groups yielded identical results.

The general conclusion can thus be stated as follows: In structured demes an individual on the average experiences its own type in greater frequency than is actually present in the deme. This causes types to interact more with similar types than with other types.

Fisher (1958), Hamilton (1963), and Trivers (1971) realized that "altruistic" traits could be selected if they were differentially directed toward fellow altruists—in other words, similar "types." Hamilton expanded this into the concept of inclusive fitness (Hamilton 1964). An application where it has been employed with great success is interactions among kin who have a high probability of similarity by descent.

This model suggests that the same evolutionary force powering kin selection also operates on all structured demes. In fact, kin selection may be considered a special case of structured demes in which the kin group equals the trait group. The mathematical equivalence of the two theories for this situation will be demonstrated below.

To study the effects of deme structure on natural selection one can apply standard fitness models to single trait groups (which can be characterized by single frequencies and densities) and calculate the weighted average of fitness over all the trait groups in the deme (D. S. Wilson 1975). If the fitness models are linear functions of frequency, one can simply substitute the subjective frequencies for the true frequencies to obtain the same result. This is the subject of Sections II and III of this paper. There it will be shown that a high variance in the composition of trait groups (hereafter referred to as "trait-group variation") promotes the evolution of altruistic and other group-advantageous traits. It is thus of interest to investigate the determinants of trait-group variation and the amount to be expected in nature.

Determinants of Trait-Group Variation

As a null hypothesis one might expect random trait-group variation, that is, types binomially distributed into trait groups. Even this can lead to significant

departures from traditional selection models (D. S. Wilson 1975 and Sections II and III). However, I hope to be able to show below that most trait-group variation in nature is greater than random.

Trait-group variation can be influenced by three separate processes: (1) differential mixing during dispersal, (2) reproduction inside trait groups, and (3) differential selection inside trait groups. I am interested in the trait-group variation prior to selection and so will only consider the first two.

Differential mixing: the environment as a segregator.—The technique of electrophoresis exploits the fact that molecules differing in charge migrate differentially in an electric field. The molecules react differently to their "environment" and hence segregate spatially. In the exact same fashion it is not hard to imagine that plankton differing in density, size, and shape might separate spatially in langmuir spirals (Hutchinson 1967); that zooplankton differing in vertical migration patterns should show considerable horizontal segregation (Miller 1970; Hardy 1956); or that, in general, morphological, behavioral, and temporal differences between types may be translated by interactions with the environment into spatial differences. Of course one does not expect anything resembling "banding" patterns because environmental forces are rarely applied in unidirectional fields. Rather, the two types will become increasingly patchy with respect to each other.

As an example, consider two types of insects, in frequency a and b, dispersing over the landscape (seeds or pelagic larvae would do as well). Divide the landscape into a grid of T equal-sized trait groups and let each trait group ($i = 1$, $2, 3, \ldots, T$) have a probability of being colonized by A types ($P_{A,i}$) and a probability of being colonized by B types ($P_{B,i}$). Then the two types will be distributed randomly with respect to each other only if

$$P_{A,i}/a = P_{B,i}/b \qquad (2)$$

for all i's (for the types to be distributed randomly with respect to space there is a further restriction that all $P_A/a = P_B/b$). Only then will trait groups of equal overall density possess a binomial distribution in their composition of types. This is a stringent requirement. Clearly it will be violated by almost any differences among types such as size, surface area, speed of flight, phototaxis, or altitudinal level sought.

More important, all insects are to varying degrees moved about involuntarily by the wind (Johnson 1969). Wind patterns are capricious, so the whole pattern of probabilities can be expected to shift from day to day and hour to hour. Thus, even if the types are structurally and behaviorally identical and all $P_{A,i}/a = P_{B,i}/b$ at any given time, if the types disperse at different average times the conditions for randomness will again be violated. It is characteristic of insects that the fittest types develop faster and emerge earlier. It is thus not unreasonable to expect differences in fitness itself to translate into partial spatial segregation of types.

The two types may have identical niche requirements, and no specific adaptations for spatial segregation need exist. Almost any differences may
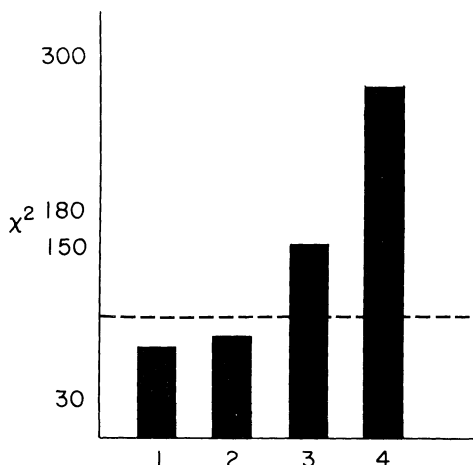
FIG. 2.—χ^2 values of bead experiments. For all bars that reach above the dashed line (the 95 percentile) the types vary nonrandomly with respect to each other.

cause spatial segregation. The concept of two types distributed randomly with respect to each other should be thought of not as a null hypothesis but as a limiting case, met only when the types are identical in every way. This was well known to the early experimenters in probability theory, who found it almost impossible to come up with a "true" coin or die (Feller 1950).

To clarify these ideas further, the following experiment was conducted. Color-coded plastic beads were obtained in the following size categories: 10–20 (two colors), 20–30, and 30–40 μm. Two "types" were selected, mixed together thoroughly in water, and plated onto a millipore filter (the "landscape"). The filter was cleared in permount and mounted on a microscope slide. The composition of 25 microscope fields (the "trait groups") was counted and subjected to a χ^2 contingency test to see if the types varied randomly with respect to each other. The results are displayed in figure 2. Each bar in the histogram represents three replicates.

The first bar portrays two types (10–20 μm), identical in every respect except color. They were poured into the millipore apparatus and plated directly onto the filter. As would be expected, they vary randomly with respect to each other ($\chi^2 = 71.85$, df = 72, $P < .5$).

The second bar portrays the same two identical types, mixed together and drawn down until 5 ml remained in the millipore apparatus. A stream of air (the "environment") was played arbitrarily over the surface through a pipette for 30 s and the remainder of the water then drawn off. This caused considerable spatial patchiness, but the two types were still distributed randomly with respect to each other ($\chi^2 = 80.73$, $.5 > P > .1$).

The third and fourth bars represent types that differ in their size (10–20, 20–30 and 10–20, 30–40 μm, respectively), subjected to the same conditions as the second bar. The different types react differently to the environmental
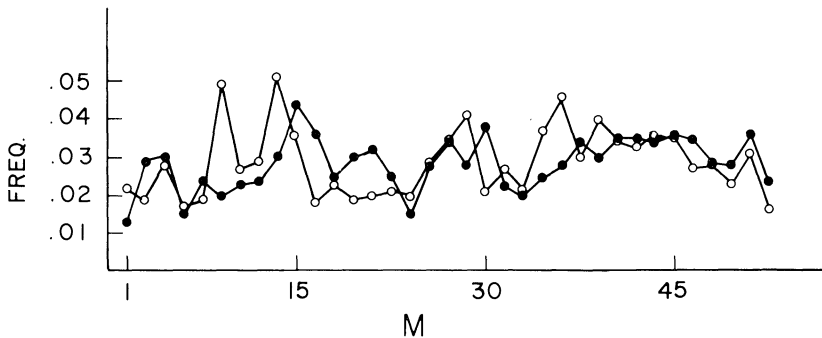
FIG. 3.—Distribution of calanoid copepods (solid circle) and cladocerans (open circle) over a 60-m horizontal transect taken in Gull Lake, Michigan.

forces and are no longer distributed randomly with respect to each other (χ^2 = 151.44 and 280.81, respectively, $P \ll .005$). The greater the difference between types, the greater the segregation.

While the bead experiments are artificial, they illustrate that high trait-group variation can form entirely through differential interactions with their environment without any specific adaptations for segregation, and there is no reason to expect live organisms to behave differently. An example from natural populations is given in figure 3, showing the horizontal distributions of calanoid copepods (closed circles) and cladocerans (open circles) over a 60-m surface transect from Gull Lake, Michigan. The data were taken with a continuous plankton recorder designed to sample small-scale distributions (< 1 m) (D. S. Wilson, unpublished). The scale is so small that the spatial distributions are probably due entirely to differential mixing, that is, differential mortality and reproduction can be excluded.

Both distributions are highly patchy through space. In relation to each other they show no obvious pattern—sometimes the peaks of abundance coincide and sometimes they do not. But that is just the point—two patchy distributions superimposed arbitrarily are patchy with respect to each other, and the variation in frequency is very much greater than random (χ^2 = 148, df = 34, $P \ll .001$).

In summary, if two types differ in morphology, behavior, or temporal activity during dispersal they cannot be expected to be distributed randomly with respect to each other. These differences will be translated by interactions with the environment into spatial differences. This is a form of segregation that requires no specific adaptations. It is an automatic consequence of differences between types. Consequently, due to this process alone trait-group variation should be at least random, usually greater.

Reproduction inside the trait group.—In this section I hope to (1) relate structured demes to kin-selection theory, and (2) show that kin selection has a significant effect (as measured by departure from random distributions) even when considerable mixing occurs among offspring of different parents.

To clarify the relation between kin selection and structured demes, consider

a sexual haploid population containing two alleles, A and B, in proportions a and b (different letters used to preserve terminology). Mating occurs at random, so the frequency of A-A, A-B, and B-B matings are a^2, $2ab$, and b^2, respectively, and

$a^2/(a^2 + ab)$ = the proportion of A offspring resulting from A-A matings,

$ab/(a^2 + ab)$ = the proportion of A offspring from A-B matings,

$b^2/(b^2 + ab)$ = the proportion of B offspring from B-B matings,

$ab/(b^2 + ab)$ = the proportion of B offspring from A-B matings.

Each female has a clutch of offspring (size N) that remains isolated from other clutches. Because interactions are restricted to within a clutch, each constitutes a trait group, and each trait group is composed entirely of siblings.

The subjective frequencies for this situation may be calculated as follows: The clutches from A-A and B-B matings will have frequencies of $a = 1$ and $a = 0$, respectively. The clutches from A-B matings will have mean frequencies of $a = b = .5$ and a binomially distributed variance of $\sigma^2 = ab/N = .25/N$. Thus the average A offspring from A-B matings will experience a subjective frequency of $a_A = .5 + .5/N$, while B offspring from A-B matings will experience a subjective frequency of $a_B = .5 - .5/N$ (eq. [1]). The subjective frequencies for all offspring from all matings are then

$$a_A = \frac{a^2(1) + ab(.5 + .5/N)}{a^2 + ab} \tag{3}$$

$$a_B = \frac{b^2(0) + ab(.5 - .5/N)}{b^2 + ab} = \frac{ab(.5 - .5/N)}{b^2 + ab}. \tag{4}$$

These are the characteristic subjective frequencies when interactions occur exclusively among siblings. In Section II, I show that the use of these values in standard donor-recipient models leads to the same conclusions as do kin selection models—namely, that the cost to the donor must be less than one-half the gain to the recipient.

In short, the fundamental process involves the existence of trait groups and the variation in the frequency of types that occurs between them. Spatial proximity of kin is one way to generate this variation, but it is not the only way. See Hamilton (1975) for another statement of this view.

If clutches from different parents mingle with one another, the variation between trait groups is diminished but not destroyed. This can best be seen by returning to an asexual population in which A and B represent two adult genotypes, in overall frequencies a and b, distributed N to a trait group. If they are distributed binomially, then the variance in frequency will be ab/N. Each female now has e offspring, which remain in the trait group. Thus each trait group consists of the intermingling offspring of N parents. The mean and variance of offspring remain the same as for the parents, but had they themselves been distributed binomially into the trait groups their variance would have been ab/eN. The effect of reproduction inside the trait groups is to raise trait-group variation by a factor e over the random case.

The structured-deme concept thus endows kin selection with a greater generality, for to date kin selection has been invoked mainly when the kin remain highly clustered in space (e.g., E. O. Wilson 1975). For example, if sibling insect larvae from a clutch of eggs are seen to mix freely with other clutches in their vicinity, it seems as if the conditions for kin selection have been violated, until one realizes that this mixing is trivial compared to that which occurs during a later dispersal state. The kin-selection process thus operates in trait groups larger than the kin group within which complete mixing occurs. Theoretically, this applies to all organisms that do not disperse immediately upon release from the parent.

Evidence from natural populations: an operational dilemma.—The previous two sections provide reasons to believe that high trait-group variation—one of the ingredients for the selection of group-advantageous traits—may be commonly found in nature, and when types are sampled over small spatial areas in nature a large variation in their frequencies is often found. But here an operational dilemma arises: Is this variation due to differential mixing and multiplication inside the trait group, or is it the product of selection due to differences in microhabitat (i.e., the multiple-niche hypothesis), a form of variation not applicable to the process investigated here? Both alternatives are probably pervasive in nature, and it is difficult to satisfactorily determine in any single case which is operating. Furthermore, it is sometimes difficult to determine if the spatial scale we are studying corresponds to demes and trait groups. Until studies are conducted in which the multiple-niche hypothesis can be excluded and trait groups defined, empirical data must be held suspect.

With this in mind, tables 1–4 present some possible examples of trait-group variation in nature for fresh water benthos, ground arthropods, marine zoo-plankton, tree-hole insects, and a population of house mice (see tables for references and more information). The numbers are the difference between the subjective and the average frequency $(a_A - a)$. The calculations included variations in density by necessity. The sampling units are quite small, although again there is no way of knowing whether they approximate the size of a trait group or if all sampling units fall within a single "deme."

Studies that found a correlation between type frequency and some habitat variable such as moisture (Allard et al. 1972) or vertical position in the inter-tidal zone (Koehn et al. 1973) are not included. There are thus no known microhabitat differences among trait groups.

If these values are at all representative, then trait-group variation in nature is often considerable. The average difference between the subjective and the true frequencies is .09 for the benthos study, .21 for the ground arthropods, and .25 for the zooplankton. Unfortunately, only one study is included here on trait-group variation in allele frequency (Selander 1970), which nevertheless shows high values (average difference = .16). (Several of the interspecific types shown are congeneric, however, and may be expected to compete actively.) In addition, Schaal (1974) investigated allele frequency variation in the annual herb *Liatris cylindracea*. Gene frequency differences of as much as .2 existed between neighboring 3-m² quadrats, with no known correlations with en-

TABLE 1

Six Species of Chironomidae

A	B					
	1	2	3	4	5	6
1....................	...	.05	.34	.02	.07	.14
2....................	.02	...	.02	.05	.02	.01
3....................	.15	.07	...	.12	.20	.16
4....................	.01	.06	.03	...	0	.01
5....................	.02	.05	.34	.01	...	.12
6....................	.12	.06	.17	.04	.12	...

Source.—Paterson and Fernando (1971).

Note.—$(a_A - a)$ within an area of 2 m². Nine subsamples were taken, each 6 × 6 inches; 1 = *Chironomous digitatus*, 2 = *C. tenvicandatus* Malloch, 3 = *Glyptotendipes barbipes* Staeger, 4 = *Polypedilum simulans* Townes, 5 = *Tanytarsus* (T.) sp., 6 = *T.* (*Cladotanytarsus*) sp. A.

TABLE 2

Seven Species of Ground Arthropods from a Beech Grove

A	B						
	1	2	3	4	5	6	7
1.	...	.16	.22	.12	0	.18	.15
2.	.10	...	.07	.17	.03	.09	.04
3.	.10	.40	...	.18	.03	.08	.09
4.	.10	.43	.18	...	.04	.14	.07
5.	.9	.24	.27	.51	...	.25	.46
6.	.11	.41	.14	.57	.21	...	.08
7.	.37	.52	.33	.85	.08	.20	...

Source.—Lloyd (1967).

Note.—$(a_A - a)$. Subsamples were 0.866 ft² in area; 1 = a woodlouse, *Trichoniscus pusillus provisorius*; 2 = an oribatid mite, *Damaeus onustus*; 3 = an oniscomorph millipede, *Glomeris marginata*; 4 = a pseudoscorpion, *Chthonius ischnocheles*; 5 = a small agelenid spider, *Hahnia helveola*; 6 = a centipede, *Lithohius crassipes*; and 7 = a staphylinid beetle, *Gyrohypnus myrmecophitus*.

vironmental parameters. M. Cahn (personal communication) divided a meter square into 10-cm² quadrats and found intense spatial segregation in the distribution of several alleles in clover. Again, there were no known correlations with environmental parameters.

Finally, D. Levin (1974), in a nearest-neighbor analysis, showed significant segregation between pins and thrums in *Hedyotis nigricans*. These studies all show that high trait-group variation can form among types, even at the allelic level.

It is now widely accepted that if altruistic traits are directed preferentially toward fellow altruists, they can be favored by natural selection. I have attempted to show that this criterion is met, to varying degrees, whenever a variation in the frequency of types exists within the deme.

TABLE 3

Ten Species of Zooplankton

					B					
A	1	2	3	4	5	6	7	8	9	10
1.	...	.49	.29	.61	.07	.25	.26	.10	.14	.17
2.	.09	...	.04	.22	.06	.11	.07	.12	.17	.44
3.	.43	.08	...	.37	.43	.23	.40	.53	.44	.13
4.	.20	.35	.11	...	.13	.25	.18	.28	.36	.21
5.	.32	.63	.46	.76	...	.49	.38	.47	.32	.22
6.	.18	.39	.17	.59	.14	...	.17	.01	.24	.08
7.	.26	.16	.30	.63	.14	.26	...	.24	.08	0
8.	.14	.61	.25	.67	.18	.12	.23	...	.19	.09
9.	.09	.50	.17	.56	.05	.13	.03	.10	...	.02
10.	.05	.34	.04	.72	.03	.05	.01	.05	.06	...

Source.—Hardy (1955).
Note.—$(a_A - a)$. The subsamples are 33 segments of an 11-mile horizontal transect.
1 = *Euphausia superba* adults; 2 = *E. superba* larvae; 3 = *Thysanoessa* post larvae;
4 = *Salpa fusiformis*; 5 = young copepods other than *Calanus*; 6 = young *Calanus*;
7 = *Calanus propinquus*; 8 = *Calanus acutus*; 9 = *Metridia gerlachei*; 10 = *Oithona* spp.

TABLE 4

More Subjective Minus True Frequencies

$A \times B$	$a_A - a$	$b_B - b$
1. MM × MS	.25	.05
2. MM × SS	.34	.12
3. MS × SS	.09	.11
4. *Metriocnemus* × *Dasyhelea*	.12	.17

Source.—Kitching (1969); Selander (1970).
Note.—1–3 = three genotypes for the Esterase-3 locus in house mice (*Mus musculus*) living in a large barn. The subdivisions are 5-ft^2 quadrats. 4 = two midge species, *Metriocnemus martinii* and *Dasyhelea* (*dufourci*?) living in tree holes; 19 tree holes were inspected.

For this to occur it is necessary that the set of organisms an individual interacts with (its trait group) be smaller than the deme. This allows the occurrence of frequency variation. Given this, the extent of frequency variation is governed by a multitude of factors. At the very least one must expect types to be binomially distributed into trait groups. This is a minimum variation, and most trait-group variation can be expected to be considerably larger. Data from natural populations indicate that trait-group variation can be quite high, even at the allelic level.

II. STRUCTURED DEMES AND THE EVOLUTION OF ALTRUISTIC TRAITS

In Section I of this paper, I argued that most demes are "structured"; that is, they are not internally homogeneous but are composed of ecological subunits, or "trait groups." This creates a spatial variation in the frequency of types

(alleles, genotypes, species) within the deme, which in turn allows the operation of an evolutionary force similar to kin selection. In fact, kin selection is part of a more general process.

Here I explore some of the implications of this process for the selection of traits, showing that under certain circumstances a form of group selection can occur. All the following models consider selection between two types, A and B. These can be conspecific individuals bearing different heritable traits or members of two species competing with each other. I do not frame my arguments in the traditional single-locus model for three reasons: (1) It is inapplicable to haploid and two-species competition; (2) it is inapplicable to traits with a more complex genetic basis; and (3) with the exception of heterosis, heterozygosity does not fundamentally alter the outcome of single-locus selection models (Lewontin 1974) and is thus a needless mathematical incumbrance. This is not to say that a consideration of heterozygosity would not be interesting—it is simply not the format for a presentation of the general concept.

A Graphical Approach

Consider two types, A and B, present in an unstructured deme at frequencies a and b ($= 1 - a$). Let $f_A(a, N)$ and $f_B(a, N)$ be the per capita (individual) fitnesses of the A and B types, respectively, which are functions of both density (N) and frequency (a). By holding N constant we can explore the relationship between frequency and fitness. Figure 4 shows linearizations of some possible fitness functions.

In figure 4A the fitness is frequency independent, and the B type (dotted line) is selected for. This is the kind of trait used in most population genetics models, which simply assign constant fitness values to the two types. It is probably less common in nature than was once thought (Ayala and Campbell 1974; Kojima 1971).

Figure 4B and C show the standard concept of frequency-dependent selection, leading either to a stable equilibrium (B) or selection for the most abundant type (C) (Ayala and Campbell 1974).

Figures 4D and E show a different sort of frequency-dependent selection, seldom mentioned because they do not seem to lead to any interesting conclusions. The B type is selected for in both cases, just as in 4A. (The lines need not be parallel.) However, in 4E the population has, in the process, reduced its own fitness. This is the dilemma of Darwinian selection around which the controversy of group selection has centered. A decrease in fitness can be selected for as long as it increases relative fitness. Darwinian selection is not sensitive to the slopes of the fitness functions—only to their positions relative to each other.

How prevalent in nature are the fitness functions represented by 4E? Some believe they are extremely common—that if nature were truly operating according to Darwinian selection it would destroy itself by this route. They thus see the necessity for a major alteration in evolutionary theory (Wynne-Edwards 1962). Others (almost everyone else) regard it as rare, only occasionally creating instabilities and extinctions, which are of course observed.
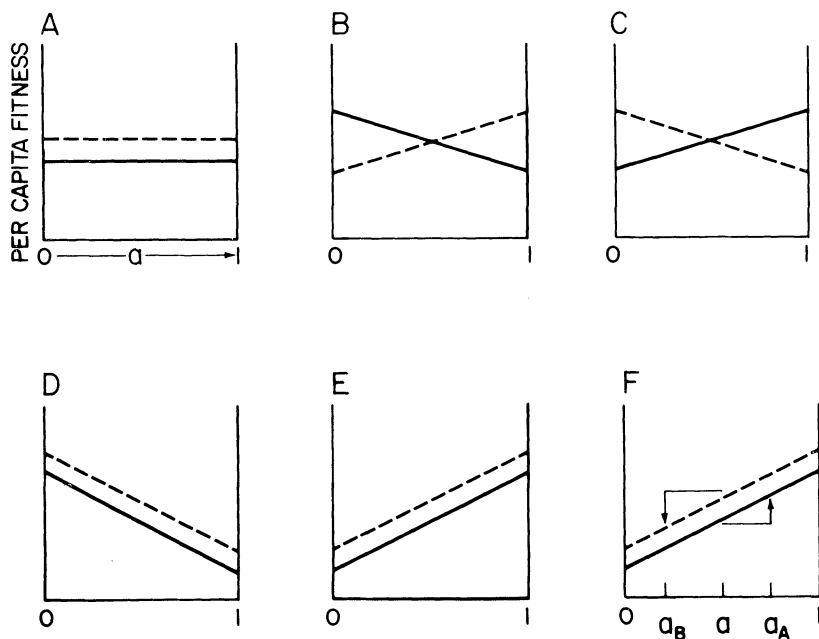
FIG. 4.—Per capita fitness of A (solid line) and B (dashed line) types as a function of frequency. Overall density (N) is held constant. A = constant fitness values for each genotype; B = frequency-dependent selection leading to stable and C = unstable equilibrium; D, E = frequency-dependent selection that does not lead to an equilibrium; E = the concept of an "altruistic" trait in which the A type increases the fitness of the population (ascending slope) but is nevertheless selected against (dashed line always above solid line); F = the effect of deme structure, which under certain circumstances can reverse the outcome of selection causing the altruist to be favored in evolution. See text for further explanation.

But regardless of the magnitude of the dilemma, it is partially resolved by the nature of structured demes, which causes each individual, on the average, to subjectively experience a higher frequency of its own type than is actually present in the deme. If a, b are the true frequencies of the types, then

$$a_A = a + \sigma^2/a = A\text{'s subjectively experienced frequency of } A,$$
$$b_A = b - \sigma^2/a = A\text{'s subjectively experienced frequency of } B,$$
$$a_B = a - \sigma^2/b = B\text{'s subjectively experienced frequency of } A,$$
$$b_B = b + \sigma^2/b = B\text{'s subjectively experienced frequency of } B,$$

where σ^2 is the variance in the frequency of types within the deme (Section I).

The effect of structured demes on the selection of traits can be seen in figure 4F. The subjective frequency and therefore per capita fitness are displaced to the right of the true frequency for the A type (a_A) and to the left for the B type (a_B, arrows). This can reverse the outcome of selection; that is, it can cause the A type to have a higher per capita fitness than the B type.

From figure 4F it can easily be seen that three things promote selection for the A type: (1) The slopes of the fitness function; i.e., the steeper the slope, the

smaller the displacement necessary to reverse the outcome of selection. (2) The vertical distance separating the fitness functions; i.e., the smaller the distance, the smaller the displacement necessary to reverse the outcome of selection. (3) The amount of trait-group variation; i.e., the larger the variation, the larger is the displacement of the subjective from the true frequencies (Section I).

Mathematical Models

The effect of deme structure on standard individual selection models is accomplished simply by substituting subjective frequencies for the true frequencies in the equations.

Warning cries and other donor-recipient relations.—Earlier (D. S. Wilson 1975) I developed a simple donor-recipient model in which $f_A(a) = f_d + (Na - 1)f_r$, $f_B(a) = Naf_r$. The per capita fitness of A includes both its role as a donor (f_d) and as a recipient to other A types $([Na - 1]f_r)$, whereas the B type serves only as a recipient. The f_r was assumed to be the same for both A and B types. Such would be the case for a warning call, in which a single A type spots a predator and warns all other A and B types in the trait group alike. The trait group in this case would be all animals within hearing range of the warner. Charnov and Krebs (1975) have produced a similar model.

Both equations form lines of identical slope, separated by the distance $(f_d - f_r)$, exactly as in figure 4E. The $f_B(a)$ equals zero at $a = 0$ because the argument in D. S. Wilson (1975) was framed in terms of fitness changes from a "baseline" level.

In unstructured demes the A type is selected for only if

$$f_A(a) > f_B(a) = f_d + (Na - 1)f_r > Naf_r$$

$$= f_d > f_r. \tag{5}$$

However, in structured demes the subjective frequencies are used:

$$f_A(a_A) > f_B(a_B) = f_d + [N(a + \sigma^2/a) - 1]f_r > N(a - \sigma^2/b)f_r$$

$$= f_d > f_r[-N(\sigma^2/a + \sigma^2/b) + 1]. \tag{6}$$

If the types are binomially distributed into trait groups, then $\sigma^2 = ab/N$, and

$$f_d > f_r[-N(ab/Na + ab/Nb) + 1]$$

$$f_d > 0. \tag{7}$$

In other words, given a binomial distribution of types into trait groups, any positive f_d will be selected for, regardless of the fitness change to the recipient (stated without proof in D. S. Wilson [1975]). This is a very different concept from traditional Darwinian selection. Notice that the process operates at all trait-group densities (N) and at all frequencies of the A type in the deme (a)—even, presumably, at the mutation frequency of the A type.

Given a greater than binomial variance in trait-group composition, truly altruistic traits that actually decrease the fitness of the donor can be selected

for. It is not known whether warning cries in nature actually fall into this category (Williams 1966; Charnov and Krebs 1975); however, if they are the altruistic traits originally envisioned by group selectionists, there should be few problems in explaining their selection.

If only a single recipient is involved in the interaction the equations must be changed to

$$f_A(a) = f_d + \frac{(Na - 1)}{(N - 1)} f_r, \qquad f_B(a) = \frac{Na}{(N - 1)} f_r. \tag{8}$$

In other words, the effect of the donor is divided among the $(N - 1)$ potential recipients. This is the type of model used in most kin-selection arguments, which relate the donor's cost to the recipient's gain. When the organisms involved are siblings this leads to the well-known conclusion that the cost must be less than one-half the gain.

Using the subjective frequencies arising from sibling interactions developed in Section I (eqq. [3, 4]):

$$f_A(a_A) > f_B(a_B)$$

$$f_d > \frac{f_r}{N - 1} \left[N(a_B - a_A) + 1 \right] \tag{9}$$

$$f_d > \frac{f_r}{N - 1} \left\{ N \left[\frac{ab(.5 - .5/N)}{b^2 + ab} - \frac{a^2 + ab(.5 + .5/N)}{a^2 + ab} \right] + 1 \right\} \tag{10}$$

$$f_d > \frac{f_r}{N - 1} \left(N \left\{ \frac{(a^2 + ab)ab(.5 - .5/N) - (b^2 + ab)[a^2 + ab(.5 + .5/N)]}{(b^2 + ab)(a^2 + ab)} \right\} + 1 \right) \tag{11}$$

$$f_d > \frac{f_r}{N - 1} \left\{ -N \left[\frac{a^3b(.5 + .5/N) + a^2b^2(1 + 1/N) + ab^3(.5 + .5/N)}{a^3b + 2a^2b^2 + ab^3} \right] + 1 \right\} \tag{12}$$

$$f_d > \frac{f_r}{N - 1} \left[-N(.5 + .5/N) + 1 \right] \tag{13}$$

$$f_d > -0.5 f_r. \tag{14}$$

This demonstrates the basic equivalence between kin selection and evolution in structured demes.

Resource notification.—Another example of seemingly "altruistic behavior" is exhibited by carrion insects and other organisms who upon finding a resource item summon their conspecifics to the scene, usually through a pheromone (Blum 1974; Borden 1974). Let P = the probability of an individual finding a resource item, M = the per capita number of offspring from a resource to which conspecifics have been summoned, and M' = the per capita number of offspring from a resource to which conspecifics have not been summoned.

Assume that the A type upon finding an item sends out a pheromone that notifies A and B types alike within a certain radius. This radius defines the trait group. The B type, however, does not summon anyone.

From simple probability theory a B type's chance of attaining a resource, either by finding it itself or through being summoned, will be $(1 - [1 - P]^{Na+1})$, while an A type's is only $(1 - [1 - P]^{Na})$. If P is small, then the probability of simultaneous discovery will be slim, and as an approximation: $P(1 + Na) =$ a B individual's chance of attaining a resource, and $PNa =$ an A individual's chance of attaining a resource.

In other words, all A types make their discoveries public, but in addition to this the B type can benefit from its private discoveries. It thus always enjoys an advantage.

Multiplying the probabilities by their respective fecundities, then, $f_A(a) = MpNa$, $f_B(a) = P(M' + MNa)$. Once again the fitness functions form parallel lines separated by the distance PM', exactly as in 4E. In unstructured demes the A type is selected for only if

$$f_A(a) > f_B(a) = MpNa > P(M' + MNa)$$

$$= 0 > M'. \qquad (15)$$

This is impossible, yet the behavior is observed.

In structured demes

$$f_A(a_A) > f_B(a_B) = MPN(a + \sigma^2/a) > PM' + MPN(a - \sigma^2/b)$$

$$= N(\sigma^2/a + \sigma^2/b) > M'/M. \qquad (16)$$

Given a binomial distribution of types into trait groups, then exactly as in the case for warning cries, $1 > M'/M$. If summoning conspecifics does not reduce per capita egg production, then $M' = M$, and the A type is selected for given a greater than binomial variance among trait groups. Such might be the case for carrion insects, in which the resource item is enormous relative to the consumer.

In some cases aggregations actually improve the per capita offspring production ($M' < M$) (Borden 1974), in which case the trait will be selected for given a binomial variance or even less. If aggregations actually decrease per capita egg production ($M' > M$), then the selection of the trait will require ever greater variances. However, notice that the size of the trait group (the effective radius of the signal) is itself subject to selection and can be adjusted until just the right number of individuals is summoned, that is, the number that can process the resource without interfering with each other ($M' = M$).

Notice that even when aggregations facilitate fecundity ($M' < M$), the trait poses a problem for individual selection, which according to this model requires that $M' < 0$. Doubtless another model could be constructed more favorable to individual selection, but this one seems quite parsimonious.

Exploitation competition: the optimal feeding rate.—This section shows that

under certain conditions a decrease in feeding rate can be selected for. Mac-Arthur (1972) considered the following paired consumer-resource relationship:

$$\frac{dR}{dt} = (r/K)R(K - R) - RNc \tag{17}$$

$$\frac{1}{N}\frac{dN}{dt} = cR - T, \tag{18}$$

where N = density of the consumer population, c = feeding rate of an individual consumer, T = energy demands (including replacement) of the individual consumer, R = the resource density, and K, r = the carrying capacity and intrinsic rate of increase of the resource (K and R expressed in "energy units" of the consumer).

Solving for $dR/dt = 0$ we get

$$R = K/r(r - cN) \tag{19}$$

$$\frac{1}{N}\frac{dN}{dt} = cR - T = (cK/r)(r - cN) - T. \tag{20}$$

MacArthur considered the equilibrium case where both dN/dt and dR/dt equal zero. However, here the resource is assumed to be at equilibrium while the consumer may be at disequilibrium. This is biologically realistic when the consumer lifetime spans many resource generations, such that the resource responds very quickly to the consumer, while the consumer responds much more slowly to the resource (zooplankton feeding on algae, fish on zooplankton, birds on multivoltine insects, cats on mice).

Notice also that equation (20), which equals the per capita fitness of the consumer, is parabolic in shape. Its optimal fitness is at an intermediate feeding rate (specifically, at $cN = r/2$), so it is capable of imprudence or overexploitation of its resources. This is the situation we desire to model.

Now consider two consumer types, A and B, in frequencies a and b. These are identical in every respect except their feeding rates, which are pc and c, respectively. Equations (19) and (20) now become:

$$R = (K/r)(r - pcNa - cNb)$$
$$= (k/r)[r - cN(pa + 1 - a)]$$
$$= (k/r)\{r - cN[1 - a(1 - p)]\} \tag{21}$$

$$\frac{1}{Na}\frac{dNa}{dt} = f_A(a) = pcR - T = pc(k/r)\{r - cN[1 - a(1 - p)]\} - T \tag{22}$$

$$\frac{1}{Nb}\frac{dNb}{dt} = f_B(a) = cR - T = c(k/r)\{r - cN[1 - a(1 - p)]\} - T. \tag{23}$$

In unstructured demes the A type is selected for only if

$$f_A(a) > f_B(a) = p > 1. \tag{24}$$

There is no optimal feeding rate. An increase is always selected for, in accordance with the traditional concept of Darwinian selection (the terms in braces in eqq. [22, 23] can never be negative, as this would imply a negative standing crop). Graphically, the fitness functions take the form similar to $4E$ but in this case they differ in their slopes.

We desire to know if an optimal feeding rate exists in structured demes. Let

$$
\begin{aligned}
y &= f_A(a_A) - f_B(a_B) \\
&= (pc(k/r)\{r - cN[1 - a_A(1 - p)]\} - T) \\
&\quad - (c(k/r)\{r - cN[1 - a_B(1 - p)]\} - T) \\
&= [pck - pc^2(k/r)N + pc^2(k/r)Na_A - p^2c^2(k/r)Na_A] \\
&\quad - [ck - c^2N(k/r) + c^2N(k/r)a_B - c^2N(k/r)a_Bp].
\end{aligned}
\tag{25}
$$

Then

$$
\frac{dy}{dp} = ck - c^2(k/r)N + c^2(k/r)Na_A - 2pc^2(k/r)Na_A + c^2N(k/r)a_B = 0
$$

$$
= c^2(k/r)N[(r/cN) - 1 + a_A + a_B - 2pa_A] = 0
\tag{26}
$$

$$
p_{opt} = [(r/cN) - 1 + a_A + a_B]/2a_A.
\tag{27}
$$

This gives the optimal value of p, that maximizes the difference between f_A and f_B. This value will be selected over all others. A lower feeding rate is selected for when $p_{opt} < 1$:

$$
\{[(r/cN) - 1 + a_A + a_B]/2a_A\} < 1
\tag{28}
$$

$$
(r/cN) < 1 + a_A - a_B
$$

$$
r < cN(1 + a_A - a_B).
\tag{29}
$$

Similarly, a higher feeding rate is selected for when

$$
p_{opt} > 1 = r > cN(1 + a_A - a_B).
\tag{30}
$$

Substituting in the values of a_B and a_A:

$$
cN\{1 - [a - (\sigma^2/b)] + [a + (\sigma^2/a)]\} > r
$$

$$
cN > \{r/[1 + (\sigma^2/a) + (\sigma^2/b)]\} \quad \text{if } p < 1
\tag{31}
$$

$$
cN < \{r/[1 + (\sigma^2/a) + (\sigma^2/b)]\} \quad \text{if } p > 1.
\tag{32}
$$

If inequality (31) holds, a decrease in feeding rate will be selected for; if (32) holds an increase is selected for. In short, an optimal feeding rate does exist at

$$
cN = r/[1 + (\sigma^2/a) + (\sigma^2/b)].
\tag{33}
$$

Notice that the optimal feeding rate is sensitive to the group consumption pressure (cN). If consumer density increases, then the individual's optimal feeding rate decreases to keep the group consumption pressure constant. Most organisms can behaviorally alter their feeding rate, and it is not unreasonable to expect the evolution of types that can "track" the optimum as their own

population density fluctuates. In such cases the consumers will never graze
the resources below a minimum level (R_{min}):

$$R = (k/r)(r - cN)$$

$$R_{min} = (k/r)(r\{r/[1 + (\sigma^2/a) + (\sigma^2/b)]\})$$

$$= K(1 - \{1/[1 + (\sigma^2/a) + (\sigma^2/b)]\}). \tag{34}$$

In a deterministic world any variance in frequency among trait groups (σ^2)
thus assures resource survival. However, in a stochastic world a very small
R_{min} might still be unstable because small populations are susceptible to
extinction through perterbation. The effect of the probability of entire trait
groups going extinct is dealt with in another section.

Interference competition.—Interference competition is exactly what the term
implies—interference with the resource-gathering activities of the types
involved. Its selective power derives from the redirection of resources during
the time and with the energy left remaining.

Interference models exist (Case and Gilpin 1974), but MacArthur's equations
can be well adapted to the concept, and in a way that emphasizes the effect of
and upon the resource. Assume the B type is exploitative, but that the A type
possesses an interference mechanism. Let:

$$f_A(a) = p_A(a)c(k/r)\{r - cN[p_A(a)a + p_B(a)b]\} - q_A(a)T \tag{35}$$

$$f_B(a) = p_B(a)c(k/r)\{r - cN[p_A(a)a + p_B(a)b]\} - q_B(a)T. \tag{36}$$

All the variables are defined as in the previous section with the exception of:

$p_B(a) = $ a function relating how A interferes with the feeding rate of B.

$p_A(a) = $ a function relating how A interferes with its own feeding rate.

$q_B(a) = $ a function relating how A interferes with the energy ex-
penditure of B. Because T includes replacement, it embraces
such things as wounding and death from fighting.

$q_A(a) = $ a function relating how A interferes with its own energy
expenditures.

The parameter space has too many dimensions to be treated adequately
here, and a more thorough treatment will be presented elsewhere (Wilson, in
preparation). However, one feature is of particular interest and is briefly
discussed below.

The simplest form of interference might be termed "nonselective," in which
the A type (the interferer) inhibits itself just as much as it inhibits the B type.
Let $p_A(a) = p_B(a) = (1 - la)$, and $q_A(a) = q_B(a) = 1$ or any constant. The l
is a constant between 0 and 1. Then

$$f_A(a) = f_B(a) = (1 - la)c(k/r)[r - cN(1 - la)] - T. \tag{37}$$

As examples, a mutant might produce a toxin without being resistant to it, or
a behavioral encounter might equally consume the time of both participants.
These traits are selected neither for nor against in individual selection models,
for while they definitely produce a change in fitness, they do not generate the

difference in fitness required for individual selection models to act. However, in structured demes (procedure exactly as in previous section):

$$f_A(a_A) = (1 - la_A)c(k/r)[r - cN(1 - la_A)] - T$$

$$f_B(a_B) = (1 - la_B)c(k/r)[r - cN(1 - la_B)] - T \tag{38}$$

$$y = f_A(a_A) - f_B(a_B)$$

$$= [c^2N(k/r)la_A - la_A ck + la_A c^2N(k/r) - l^2a_A{}^2c^2N(k/r)]$$

$$- [c^2N(k/r)la_B - la_B ck + la_B c^2N(k/r) - l^2a_B{}^2c^2(k/r)N]$$

$$= 2c^2N(k/r)l(a_A - a_B) - ckl(a_A - a_B) - l^2c^2N(k/r)(a_A{}^2 - a_B{}^2) \tag{39}$$

$$\frac{dy}{dl} = 2c^2N(k/r)l(a_A - a_B) - ck(a_A - a_B) - 2lc^2N(k/r)(a_A{}^2 - a_B{}^2) = 0$$

$$2l(a_A - a_B)(a_A + a_B) = [2 - (r/cN)](a_A - a_B)$$

$$l_{opt} = [2 - (r/cN)]/2(a_A + a_B). \tag{40}$$

Nonselective interference is selected for when $l_{opt} > 0$. The term $(a_A + a_B)$ is always positive, so

$$2 - (r/cN) > 0, \tag{41}$$

$$cN > (r/2). \tag{42}$$

In sum, nonselective interference will be selected for whenever the group consumption rate (cN) exceeds $r/2$. Recall that $cN = r/2$ is the peak of the parabola representing per capita fitness in equation (30)—that is, maximum prudence on the part of the consumer.

What has happened to produce this result, that makes interference so much more powerful than exploitation in selecting for prudence? Graphically, through nonselective interference the A type has erased the difference between the fitness function in $1E$, which is instead represented by the same ascending line, as shown in figure 5. Now any amount of trait-group variation will cause selection of the trait. This is discussed further in Section III.

A more detailed consideration of interference competition will be given elsewhere. It suffices to say that this model predicts a powerful evolutionary force promoting prudence through interference in structured demes.

Persistence, or the probability of extinction.—Most group-selection models consider traits that increase a type's relative fitness on the one hand, but also the probability of group extinction on the other. Selection is thus viewed as a balance between relative fitness favoring one type within groups and differential extinction favoring the other type among groups (see E. O. Wilson 1973 for a review, also Gilpin 1975).

So far it has been shown that group-advantageous traits can be selected for in structured demes even without differential trait-group extinction. The inclusion of extinction further strengthens the process. Obviously, if $E(a)$ is the probability of an entire trait group going extinct, as a function of type
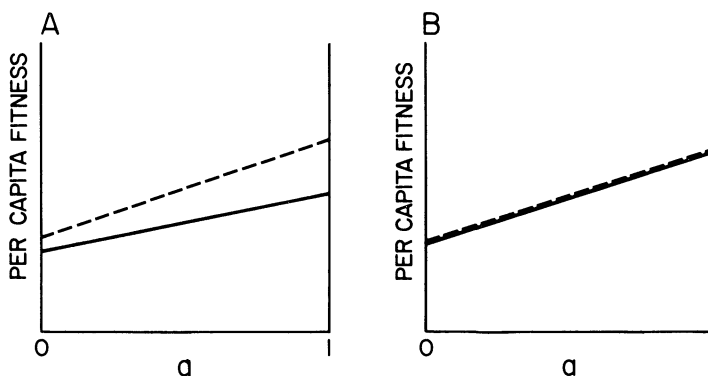
Fig. 5.—The evolution of prudence. Solid line is per capita fitness of A, dashed the per capita fitness of B. In exploitation competition ($2A$) a barrier exists for the evolution of a lower feeding rate that must be overcome by trait-group variation. However, no such barrier exists for nonselective interference ($2B$), which is therefore selected for given any trait-group variation. See text for further explanation.

frequency, then this can be multiplied with any of the fitness functions of the previous sections to obtain new functions,

$$f'_A(a_A) = E(a_A)f_A(a_A), \qquad f'_B(a_B) = E(a_B)f_B(a_B), \tag{43}$$

promoting the most persistent type.

Discussion

Recall figure $4E$, which I consider to be a precise definition of an altruistic trait. Two questions may be asked about it: (1) Can it be selected for? (2) If so, what traits in nature fall under this category? If the answer to the first question is no, or hardly ever, then the answer to the second is automatically given, and it remains the task of scientists to show that traits appearing to fall into this category actually belong to another. This is the spirit in which group selection has been debated over the past decade. Given the weakness of Wynne-Edward's (1962) mechanism, almost any alternative explanation seemed superior.

I believe that the concept of structured demes provides a strong mechanism for the selection of altruistic and group-advantageous traits. The ingredients for its operation—trait groups that vary in their composition of types—are widely found in nature. Conceptually it is a simple generalization of kin selection, and many of the conclusions hold given only random distributions of types into trait groups. Most importantly, the process is fully compatible with the principle of individual selection and does not pose any theoretical dilemmas.

If this is true, and the answer to the first question is yes, or quite often, then the second question must be approached in an entirely new spirit. It no longer suffices to show how a trait might be favored by Darwinian selection. That is only the first step in the debate, the construction of alternative hypotheses, then to be tested.

At present the only honest answer to be given to the second question is that we do not know what traits in nature can be termed altruistic. Very few traits are understood well enough to even ascribe a fitness function to them. We do not know if warning cries actually entail a risk to the donor. Above all, even assuming that prudence is possible among consumers, we do not know the extent to which they find themselves in situations where they can exercise it. Considering the millions of insects on a single bird's territory, it is entirely possible that birds are never in a position to overexploit. For this reason I have avoided "supporting" my models with data from nature. All such support is useless until fitness functions can be rigorously determined.

While the magic, truly decisive experimental test is a myth (Lewontin 1974), this theory is testable, and there is nothing preventing it from moving beyond the speculatory stage. In addition to rigorous field studies, structured and unstructured demes may easily be set up in the laboratory. If natural populations are brought in and run in parallel, one in structured and the other in unstructured demes, a divergence in their traits should become evident after several generations. If the unstructured populations diverge radically from both the structured lines and natural populations, this would be evidence that structured demes are a closer approximation to natural situations.

III. ALTRUISM, SELFISHNESS, AND GROUP-ADVANTAGEOUS TRAITS

In discussions on group selection a dichotomy is usually constructed between "selfish" behaviors that benefit the individual performing them and "altruistic" behaviors that benefit the group to the detriment of the individual performing them. The inference is that if most behaviors are selfish then selection is operating mainly on the individual level (e.g., E. O. Wilson 1975; Alexander 1974; Williams 1966).

I believe that this is a false dichotomy and that the focus of the group-selection controversy has in fact been misplaced. The fundamental question is, Does natural selection maximize population fitness? Whether it does so through selfish or altruistic traits is of secondary importance.

In the previous two sections I argued that altruistic traits could commonly be selected for in structured demes. Here I discuss the relationship between altruism, selfishness, and population fitness in a more general sense.

Fitness States and Apportionment Ratios

Recall figures $4D$ and E. When the fitness functions have a positive slope $(4E)$, the A trait has been called "group advantageous," for it increases the average per capita fitness within the group. If the functions have a negative slope $(4D)$, they may be termed "group detrimental." If A's per capita fitness (f_A) lies above B's (f_B), the trait has been termed "selfish," for it is selected for in individual selection models at all times. Of course, a selfish trait may be either group advantageous or detrimental, as is widely appreciated.

Finally, if A's per capita fitness lies below B's, the trait has been termed

TABLE 5

DEFINITIONS

Selfishness	$P_A/P_B > 1$
Altruism	$P_A/P_B < 1$
Group advantageous	$S(a)$ an increasing function of a
	$= S(a_B)/S(a_A) < 1$
Group detrimental	$S(a)$ a decreasing function of a
	$= S(a_B)/S(a_A) > 1$

"altruistic," but only when it is simultaneously group advantageous. Otherwise it is simply maladaptive. Maladaptive traits are hereafter ignored because they are selected against under all circumstances.

These relationships may be expressed mathematically using the following terms: $S(a)$ = the state of the fitness component, as a function of the frequency of A; and P_A, P_B = the apportionment of the fitness component to the A and B types, respectively.

The $S(a)$ describes the effect of the trait on the fitness component (standing crop of resource, immunity to predators, etc.), while P_A and P_B describe how it is apportioned among the types. As an example, a warning cry increases the immunity of the population to predation ($S[a]$ has a positive slope), but this immunity is apportioned preferentially to nonwarners ($P_B > P_A$).

The per capita fitnesses are the product of the fitness state and the apportionments:

$$f_A = P_A S(a), \qquad f_B = P_B S(a). \tag{44}$$

The definitions of altruism, selfishness, and group-advantageous traits using these terms are summarized in table 5. The purpose of this model is to express some of the essential features of the more specific (and complex) models of Section II in a simpler form. For example, as always the A type is selected for only if it has a higher per capita fitness, and in unstructured demes:

$$f_A > f_B$$

$$P_A S(a) > P_B S(a)$$

$$P_A/P_B > 1. \tag{45}$$

This leads to the conclusion, already arrived at many times, that individual selection is unresponsive to the state of the fitness component, only to the apportionment ratio. Only selfish behaviors are selected for, regardless of whether they are group advantageous or detrimental.

However, in structured demes the subjective frequencies always bear the relation $a_A > a > a_B$, and

$$f_A > f_B$$

$$P_A S(a_A) > P_B S(a_B)$$

$$P_A/P_B > S(a_B)/S(a_A). \tag{46}$$

In structured demes evolution is sensitive to both the apportionment ratio (P_A/P_B) and the fitness-state ratio $(S[a_B]/S[a_A])$.

Social versus Nonsocial Behaviors

Before continuing it is necessary to explore the relation between these two ratios. Consider a population in which a decrease in feeding rate is group advantageous (prudence). An A type with a lower feeding rate causes a fitness-state ratio < 1. (This follows directly from the definition of a group-advantageous trait [table 5]. If $S[a]$ is an increasing function of a and $a_A > a_B$, then $S[a_B]/S[a_A] < 1$.) However, the same trait simultaneously causes a decrease in the apportionment ratio (an altruistic trait). The two are inseparable. In essence the A type is saying, "I am going to feed less, regardless of how you feed." Models such as this give the impression that most group-advantageous traits are by nature altruistic.

However, now consider the same population with an A type that develops an interference mechanism. The simplest form of interference can be termed "nonselective," in which the A type interferes with itself as much as with the B type. (This is treated in detail in Section II. Insofar as individual selection models require the trait to interfere more with the B than with the A type to be selected for, this is hardly a restrictive assumption.) In essence the A type is now saying, "I am going to feed less, and so are you!" The same decrease in the fitness-state ratio results (a group-advantageous trait) but without the corresponding decrease in the apportionment ratio.

The point is that the two ratios are tightly coupled only when the A type does not directly influence the behavior of conspecifics, interacting only with the resource instead ("nonsocial" behavior). In such cases, group-advantageous traits are often by nature altruistic because the A type has only its own behavior to modify. But when the A type modifies the behavior of conspecifics as well as itself ("social" behavior), the two ratios become decoupled and group-advantageous traits (in this case a lower feeding rate) can as easily be altruistic (A decreases its own feeding rate), selfish (A decreases B's feeding rate), or involve no change in the apportionment ratio at all (both feeding rates decreased equally). As a second example, a decrease in population size may be accomplished through voluntary exit or expulsion.

In summary, when social as well as nonsocial behaviors are considered, there is usually no necessary relation between group advantageousness and altruism. Put another way, a given end can usually be accomplished through both selfish and altruistic means. Having stated the rule, however, it is necessary to mention an important exception. Some "emergency" situations are by necessity altruistic. A warning caller has no choice but to expose his presence to the predator.

The Evolution of Group-advantageous Traits

Return now to the conditions for selection of the A type in structured demes:

$$P_A/P_B > S(a_B)/S(a_A). \tag{47}$$

Consider a genetically homogenous population (one type), in which case both ratios equal one by definition. Mutant traits will cause deviations from one in both ratios. If the traits are social, the deviations in the two ratios will be relatively independent of one another. Several conclusions may be drawn:

1. If the trait is group advantageous, then the fitness-state ratio is less than one and deviations in the apportionment ratio in either direction—selfish or altruistic—will be selected for, although very altruistic traits are still selected against. Conversely, if the trait is group detrimental, then the fitness-state ratio is greater than one and deviations in the apportionment ratio in either direction—selfish or altruistic—are selected against, although very selfish traits are still favored. In short, a barrier operates against the evolution of group-detrimental traits that does not exist for group-advantageous traits. This can be considered an evolutionary force on the population level. Its power is unknown and must be decided by future study. It cannot completely eliminate group-detrimental traits, for the barrier is only partial. However, insofar as the same "tools" of interference are available to both group-advantageous and detrimental types, the extra advantage bestowed upon the former, even if slight, could cause it to predominate in most cases.

2. Let us assume that a group-detrimental trait arises that is so strongly selfish it is nevertheless selected for. The old group-advantageous type goes extinct, and the fitness state is lowered. The population is again homogenous and the ratios set back to one. The tendency of the system will still be toward higher fitness states but against a new background. The old trait will never again be selected for, but, rather, a new trait that both raises the fitness state and is resistant to the group-detrimental type. In other words, it is possible that social structures not only tend toward higher fitness states but also possess a continuously evolving "immunity" against previously evolved, group-detrimental types.

3. The greater the difference between the apportionment ratio and the fitness-state ratio, the more intensely the trait is favored. Therefore, a given increase in fitness state arrived at via a selfish mechanism (high P_A/P_B) will be selected over an identical increase in fitness state arrived at via an altruistic mechanism (low P_A/P_B). In other words, altruistic traits will exist only when the same result, in terms of group advantageousness, cannot be created selfishly. A preponderance of selfish traits is not an argument against group selection. It is expected.

At the risk of anthropomorphizing, an analogy may be constructed with human societies. Very few vital functions in society are performed altruistically. They are made individually advantageous. Superficially, one sees individuals relentlessly pursuing their own welfare. Less conspicuous are the forces causing them to simultaneously pursue the welfare of the group.

IV. SUMMARY

1. Most organisms interact with a set of neighbors smaller than the deme (its trait group). Demes therefore are not only a population of individuals but also a population of groups (structured demes).

2. Trait groups vary in their composition. The minimum variance to be expected is that arising from a binomial distribution. Most populations have a higher variance than this due to (a) differential interactions with the environment and (b) the effects of reproduction inside the trait groups.

3. As a consequence of this variation, an individual on the average experiences its own "type" in a greater frequency than actually exists in the deme. Its behaviors are therefore directed differentially toward fellow types, and this is the fundamental requirement for the evolution of altruism.

4. Models are presented for warning cries and other donor-recipient relations, resource notification, the evolution of prudence in exploitation and interference competition, and the effect of differential trait-group extinction. In all cases evolution in structured demes differs from traditional individual-selection models. Individual selection corresponds to the case where there is zero variance among trait groups, that is, complete homogeneity.

5. The "threshold" variance permitting the evolution of altruism (negative fitness change to the donor) is that arising from a binomial distribution. As this is the minimum to be expected in nature, this theory predicts that at least weakly altruistic behavior should be a common occurrence (but see [9]).

6. If a population is overexploiting its resource, a decrease in feeding rate through interference may be selected for given any trait-group variation.

7. When trait groups are composed entirely of siblings (i.e., kin groups), the model is mathematically equivalent to kin selection.

8. As well as increasing population fitness, socia l systems may also evolve an "immunity" against group-detrimental types.

9. If a given group-advantageous effect can be accomplished through both altruistic and selfish mechanisms, the selfish mechanism will be selected. A paucity of altruistic behaviors may signify that it is usually possible to create the same result selfishly—not that altruism "cannot" be selected for.

LITERATURE CITED

Alexander, R. D. 1974. The evolution of social behavior. Annu. Rev. Ecol. Syst. 5:325–383.

Allard, R. W., G. R. Babbel, M. T. Clegg, and A. L. Kahler. 1972. Evidence for coadaptation in *Avena barbata*. Proc. Nat. Acad. Sci. 69:3043–3048.

Ayala, F. S., and C. A. Campbell. 1974. Frequency dependent selection. Annu. Rev. Ecol. Syst. 5:115–139.

Blum, M. S. 1974. Pheromonal basis of social manifestation in insects. Pages 190–199 *in* M. C. Birch, ed. Pheromones. American Elsevier, New York.

Borden, J. H. 1974. Aggregation pheromones in the Scolytidae. Pages 135–160 *in* M. C. Birch, ed. Pheromones. American Elsevier, New York.

Bradshaw, A. D. 1972. Some of the evolutionary consequences of being a plant. Evolut. Biol. 5:25–44.

Case, T. J., and M. E. Gilpin. 1974. Interference competition and niche theory. Proc. Nat. Acad. Sci. 71:3073–3077.

Charnov, E. L., and J. R. Krebs. 1975. The evolution of alarm calls: altruism or manipulation? Amer. Natur. 109:107–112.

Christiensen, F. B., and M. W. Feldman. 1975. Subdivided populations: a review of the one and two locus deterministic theory. Theoret. Pop. Biol. 7:13–39.

Ehrlich, P. R., and P. H Raven. 1969. Differentiation of populations. Science 165:1228–1232.

Feller, W. 1950. An introduction to probability theory and its applications. Wiley, New York.

Fisher, R. A. 1958. The genetical theory of natural selection. 2d ed. Dover, New York.

Gilpin, M. E. 1975. The theory of group selection in predator prey communities. Princeton University Press, Princeton, N.J.

Hamilton, W. D. 1963. The evolution of altruistic behavior. Amer. Natur. 97:354–356.

———. 1964. The genetical evolution of social behavior. I and II. J. Theoret. Biol. 7:1–52.

———. 1971. Selection of selfish and altruistic behavior in some extreme models. Pages 57–93 *in* J. F. Eisenberg and W. S. Dillon, eds. Man and beast: comparative social behavior. Smithsonian, Washington, D.C.

———. 1975. Innate social aptitudes in man, an approach from evolutionary genetics. Pages 133–157 *in* R. Fox, ed. Biosocial anthropology. Wiley, New York.

Hardy, A. C. 1955. A further example of the patchiness of plankton distributions. Pergamon, New York.

———. 1956. The open sea: its natural history. Collins, London.

Hutchinson, G. E. 1967. A treatise on limnology. Vol. 2. Introduction to lake biology and the limnoplankton. Wiley, New York.

Johnson, C. G. 1969. Migration and dispersal of insects by flight. Methuen, London.

Kitching, R. L. 1969. The fauna of tree-holes in relation to environmental factors. Ph.D. diss. Oxford University.

Koehn, R. K., F. J. Turano, and J. B. Mitton. 1973. Population genetics of marine pelecypods. II. Genetic differences in microhabitats of *Modiolus demissus*. Evolution 22:100–106.

Kojima, K. 1971. Is there a constant fitness value for a given genotype? No! Evolution 25:281–285.

Levin, D. A. 1974. Spatial segregation of pins and thrums in populations of *Hedyotis nigricans*. Evolution 28:648–656.

Levin, D. A., and H. W. Kerster. 1968. Local gene dispersal in *Phlox pilosa*. Evolution 22:130–139.

———. 1974. Gene flow in seed plants. Evolut. Biol. 7:139–220.

Levin, S. A. 1974. Dispersion and population interactions. Amer. Natur. 108:207–228.

Lewontin, R. C. 1974. The genetic basis of evolutionary change. Columbia University Press, New York.

Lloyd, M. 1967. Mean crowding. J. Anim. Ecol. 36:1–30.

MacArthur, R. H. 1972. Geographical ecology. Harper & Row, New York.

Mayr, E. 1963. Animal species and evolution. Belknap, Cambridge, Mass.

Miller, C. B. 1970. Some environmental consequences of vertical migration in marine zooplankton. Limnol. Oceanogr. 15:727–742.

Paterson, C. G., and C. H. Fernando. 1971. Studies on the spatial heterogeneity of shallow water benthos with particular reference to the Chironomidae. Can. J. Zool. 49:1013–1019.

Schaal, B. A. 1974. Isolation by distance in *Liatris cylindracea*. Nature 252:704.

Selander, R. K. 1970. Behavior and genetic variation in natural populations. Amer. Zool. 10:53–66.

Trivers, R. L. 1971. The evolution of reciprocal altruism. Quart. Rev. Biol. 46:35–57.

Williams, G. C. 1966. Adaptation and natural selection: a critique of some current evolutionary thought. Princeton University Press, Princeton, N.J.

Wilson, D. S. 1975. A theory of group selection. Proc. Nat. Acad. Sci. 72:143–146.

Wilson, E. O. 1973. Group selection and its significance for ecology. BioScience 23:631–638.

———. 1975. Sociobiology. Harvard University Press, Cambridge, Mass.

Wynne-Edwards, V. C. 1962. Animal dispersion in relation to social behavior. Oliver & Boyd, London.