Sociobiological Variations on a Mendelian Theme

M. FRANK NORMAN

1. Introduction. My articles in this volume illustrate two aspects of my work as a psychologically oriented mathematician. The present article shows elementary mathematics in the service of psychology. The other article [1] shows elementary psychology in the service of mathematics. Both papers bear on simple genetic models: deterministic models in this paper and stochastic models in the other one.

In psychology one frequently encounters facile evolutionary explanations of contemporary behaviors and their neural substrates. It is often easy to convince oneself that a certain characteristic of a currently predominant genotype permitted individuals of that genotype to have more offspring than conspecifics, thus ensuring the success of the genotype. The (modest) interest in such exercises is partially predicated on the validity of the transition between individual reproductive success ("fitness") and long-term success of the genotype. This paper shows that such transitions are not always valid.

2. Fitness and survival. According to E. O. Wilson, "Hamilton's theorem on altruism consists merely of a more general restatement of the basic axiom that genotypes increase in frequency if their relative fitness is greater" [2, pp. 415-416, italics added]. W. D. Hamilton's theory of the evolution of altruism will be considered in §4. The present section relates to the italicized proposition, which is an unusually explicit statement of a dominant theme of the literature of Evolutionary Biology.

Consider the following example. Suppose that there are three interbreeding varieties of zebras in a certain region. Call these varieties a, b, and c. They differ in probability of survival from conception to reproductive age, perhaps because of different diets. The common term for this survival probability is viability. Assuming that all varieties are equally fertile, viability is proportional to, and can thus be identified with, the expected number of offspring of a newly

1980 Mathematics Subject Classification. Primary 92A10, 92A20, 92A25.

188

conceived zebra. This number is termed fitness, and denoted w_a , w_b , w_c for the three varieties. Suppose, for definiteness, that $w_a = 1/5$, $w_b = 3/5$, $w_c = 4/5$. Obviously, the relative frequency of c will increase from generation to generation, in accordance with Wilson's proposition, and this variety will eventually displace the other two, in accordance with the notion of "survival of the fittest."

This is "obvious," but it need not be true. As in most casual evolutionary arguments, I have said nothing about the genetic structure of my zebras. This is tantamount to assuming that genetic structure is irrelevant to the course of evolution. This is a dangerous assumption, to say the least.

Suppose, for example, that the differences between a, b, and c are controlled by one genetic locus with two alleles, A_1 and A_2 , and that the genotypes of a, b, and c are $a = A_2A_2$, $b = A_1A_1$, $c = A_1A_2$. The fitnesses w_a , w_b , and w_c will henceforth be denoted w_{22} , w_{11} , and w_{12} . Make all of the standard simplifying assumptions: random mating, infinite population, and nonoverlapping generations. Let

 p_n = proportion of A_1 genes among newly conceived individuals in the *n*th generation

and $q_n = 1 - p_n$. Then the proportions of A_1A_1 , A_1A_2 , and A_2A_2 are p_n^2 , $2p_nq_n$, q_n^2 at conception (this is the Hardy-Weinberg law [3, p. 4]) and proportional to $w_{11}p_n^2$, $2w_{12}p_nq_n$, $w_{22}q_n^2$ among adults, from which it follows that

$$p_{n+1} = \frac{w_{11}p_n^2 + w_{12}p_nq_n}{w_{11}p_n^2 + 2w_{12}p_nq_n + w_{22}q_n^2}.$$
 (2.1)

Since the heterozygote, $A_1A_2 = c$, has greater fitness than either homozygote, p_n converges to an internal equilibrium value,

$$p_{\infty} = \frac{w_{12} - w_{22}}{2w_{12} - w_{11} - w_{22}} = \frac{3}{4}.$$

The asymptotic genotype proportions, derived from the Hardy-Weinberg law, are given in Table 1.

TABLE 1
Fitness and asymptotic frequency

Variety	Genotype	Fitness	Frequency
a	A_2A_2	1/5	1/16
b	A_1A_1	3/5	9/16
c	A_1A_2	4/5	6/16

Clearly "survival of the fittest" does not apply here, if this phrase is understood to imply that the fittest variety displaces all competitors. In our example, all three varieties remain in appreciable frequencies. Moreover, the second fittest variety, b, is asymptotically more numerous than the fittest variety, c. Nor is it

correct that the fittest variety always increases in frequency, as Wilson alleges. For it can be shown that, if $p_1 < p_{\infty}$, then p_n increases to $p_{\infty} = 3/4$. Hence $P(c) = 2p_n q_n$ decreases for all n sufficiently large that $p_n > \frac{1}{2}$.

It is sometimes claimed that the theory of evolution by natural selection is vacuous, since its mechanism is "survival of the fittest" and fitness is often defined as survival probability. This conundrum arises from an ambiguity in the word "survival": it is used to refer to both individuals and varieties. Our example shows that superior individual survival probability (fitness) does not guarantee superior asymptotic frequency of the corresponding variety. (For a fuller discussion of the claim that natural selection is tautological, together with a different resolution, see [4, Chapter 4].)

The notion that one may equate individual and variety survival is deeply ingrained in biology. This notion can lead to error because it ignores the subtleties of Mendelian genetics. Its persistence is perhaps partly explicable by the long delay between publication of *The Origin of Species* and the rediscovery and acceptance of Mendel's ideas. Moreover, the linkage between fitness and asymptotic prevalence is often as direct as anyone could wish. If

$$\min\{w_{11}, w_{22}\} \leq w_{12} < \max\{w_{11}, w_{22}\},\$$

then the fittest genotype (either A_1A_1 or A_2A_2) displaces all others at asymptote. However, if $\min\{w_{11}, w_{22}\} < w_{12} = \max\{w_{11}, w_{22}\}$, then the simple fitness-prevalence linkage again fails, since A_1A_2 is completely displaced by a homozygote with the same fitness.

Our findings in this section apply with equal force to the evolution of gross anatomical structures and to the evolution of neural microstructures that influence behavior. Subsequent sections have a specifically behavioral focus. Variants of the example presented in this section will be used to expose weaknesses in two popular theories of the evolution of social behavior: Maynard Smith's theory of the evolution of behavior in conflicts between animals, and Hamilton's theory of the evolution of altruism.

3. Evolutionarily stable strategies. The central notion in Maynard Smith's theory [5] is the evolutionarily stable strategy or ESS. Roughly speaking, an ESS is a frequency distribution over alternative genetically controlled behavior patterns that is resistant to incursion by small mutant or migrant groups. Stated in this way, it is possible to apply the notion directly to our three types of zebras with different dietary preferences and corresponding fitnesses $w_a = 1/5$, $w_b = 3/5$, $w_c = 4/5$. From a naturalistic, genetically naive, viewpoint, it would appear that the ESS for zebras is for them all to be of type c. For $w_c > w_a$ and $w_c > w_b$, so mutants or migrants of types a and b would be at a disadvantage in a population of individuals of type c (so the argument goes). However, the "all c" strategy is not, in fact, stable in the ordinary, dynamic, sense, if, as before, type c zebras are A_1A_2 heterozygotes. For if all zebras were of type c in the nth

generation, then, in the next generation, types a, b, and c would have frequencies 1/4, 1/4, and 1/2. Another way of putting it is that the "all c" strategy is not available to the population under natural conditions. It ignores the constraints of Mendelian segregation. Oster and Rocklin [6, p. 31] have also noted the genetic naiveté of Maynard Smith's theory.

It may reasonably be objected that the preceding paragraph slights ESS by generalizing it beyond the game-theoretic context in which it was formulated. This context will now be described. Suppose that a population consists of k behaviorally distinct varieties of a certain species, which we label $1, 2, \ldots, k$. The distinctive behaviors are displayed in certain interactions (e.g., territorial conflicts) with other members of the population. We suppose, for simplicity, that an individual has exactly one such interaction in his lifetime, and that this interaction determines his fitness. The fitness of an i that interacts with a j is W(i, j). Consequently, the expected fitness of an i that interacts with a randomly chosen opponent is $W(i, Q) = \sum_{j=1}^{k} W(i, j)Q_{j}$, where Q_{j} is the probability that the other animal is a j. Finally, if both contestants are randomly chosen, the expected fitness of the first is

$$W(P, Q) = \sum_{i=1}^{k} \sum_{j=1}^{k} P_i W(i, j) Q_j.$$

Here one should think of choosing the two contestants independently from subpopulations with respective distributions P and Q.

The definition of an ESS is formulated in terms of the expected fitness function. For any P, Q, and $\epsilon > 0$, let $P_{\bullet} = (1 - \epsilon)P + \epsilon Q$.

DEFINITION. A distribution, P, is an ESS if, for every $O \neq P$,

$$W(P, P_{\epsilon}) > W(Q, P_{\epsilon}),$$

for all sufficiently small ε .

Here Q corresponds to a mutant or migrant subpopulation attempting, unsuccessfully, to "gain a foothold" against an established population with distribution P. This formulation of the definition of ESS follows [7].

Maynard Smith proposes that distributions occurring in nature should be evolutionarily stable in this sense. This is an appealing notion, but it may be inconsistent with certain models for the genetic substrate, as we can see by reconsidering our zebras. Thus k = 3, and varieties 1, 2, and 3 correspond to our earlier $a = A_2A_2$, $b = A_1A_1$, and $c = A_1A_2$. Moreover, we assume that the fitnesses given earlier apply here, regardless of the other animal's behavior. In other words, W(1,j) = 1/5, W(2,j) = 3/5, W(3,j) = 4/5, for j = 1, 2, 3. Thus we have simply embedded our earlier example in the present game-theoretic framework. As was noted previously, the distribution that evolves is

$$P_1 = 1/16$$
, $P_2 = 9/16$, $P_3 = 6/16$.

This distribution is not an ESS. It is easy to show that the only ESS is our old friend, the distribution concentrated on the heterozygous variety, 3.

191

The assumption that fitness is independent of the other animal's behavior makes this a degenerate example of an "interaction," much less a "conflict." However, I see nothing to prevent such inconsistencies between ESS and genetics from arising in bona fide conflict situations. The gap between ESS and genetics is clearly revealed by the fact that the rationale for ESS, such as it is, applies with equal force to interspecific and intraspecific conflict, while only the latter situation imposes genetic constraints.

Our example shows that the ESS notion is incompatible with one possible specification of genetic substrate. Perhaps it is more interesting to ask whether there is any genetic substrate with which ESS is consistent. I was surprised to discover, quite recently, that the answer is "yes." The anomalies of the zebra example are traceable to the superior fitness of heterozygotes. This led me to consider a scheme that assigns intermediate fitness to heterozygotes.

Suppose that there are only two behavioral phenotypes, 1 and 2 (e.g., the "hawks" and "doves" discussed in [5]). Underlying these are three genotypes, A_1A_1 , A_2A_2 , and A_1A_2 . All A_1A_1 homozygotes, and a proportion, γ , of the heterozygotes are assumed to have phenotype 1; the remaining heterozygotes and the A_2A_2 homozygotes have phenotype 2. As in §2, let p_n be the proportion of A_1 genes in the *n*th generation, and let P_n be the corresponding distribution over the two phenotypes. Then $P_{n,1} = p_n^2 + 2\gamma p_n q_n$. Genotypes A_2A_2 , A_1A_1 , and A_1A_2 have respective fitnesses $W(2, P_n)$, $W(1, P_n)$, and $\gamma W(1, P_n) + (1-\gamma)W(2, P_n)$, and the trajectory of the system is determined by (2.1) with w_{22} , w_{11} , and w_{12} replaced by these values.

For this model, evolutionary stability and dynamic stability are synonymous. This can be seen from the following catalog of cases corresponding to different values of the parameters

$$\delta_1 = W(1, 1) - W(2, 1)$$
 and $\delta_2 = W(2, 2) - W(1, 2)$.

Proofs are omitted. It is assumed that $0 < P_{1,1} < 1$, and P_{∞} denotes the limit of P_n as $n \to \infty$.

- (a) $\delta_1 = 0$ and $\delta_2 = 0$. Then there is no ESS, and $P_{\infty} = P_1$.
- (b.1) $\delta_1 > 0$, $\delta_2 < 0$, and at least one of these inequalities is strict. Then the only ESS is $P_{\infty} = (1, 0)$.
- (b.2) $\delta_1 < 0$, $\delta_2 > 0$, and at least one of these inequalities is strict. Then the only ESS is $P_m = (0, 1)$.
- (c) $\delta_1 < 0$ and $\delta_2 < 0$. Then the only ESS is $P_{\infty} = (\alpha, 1 \alpha)$, where $\alpha = \delta_2/(\delta_1 + \delta_2)$.
 - (d) $\delta_1 > 0$ and $\delta_2 > 0$. Then both (1, 0) and (0, 1) are ESSs. Moreover

$$P_{\infty} = \begin{cases} (1,0) & \text{if } P_{1,1} > \alpha, \\ (\alpha, 1 - \alpha) & \text{if } P_{1,1} = \alpha, \\ (0, 1) & \text{if } P_{1,1} < \alpha. \end{cases}$$

This model can be generalized in an obvious way to more than two behavioral phenotypes. It remains to be seen whether the identity between evolutionary and

dynamic stability holds for the generalization. Moreover, there is no reason to believe that such models accurately represent the genetic contribution to animals' behavior in conflict situations.

In this section we have seen that it is possible to find a genetic model compatible with Maynard Smith's theory, as well as a model that is not compatible. This leaves one uneasy about most of the literature on ESS, since this literature characteristically avoids explicit consideration of genetic substrate. Two exceptions are [7] and [8], which, however, concentrate on behaviorally uninteresting haploid organisms.

4. The evolution of altruism. Altruistic behavior involves personal sacrifice for the benefit of others. In the present context, we must think of loss of personal fitness in the process of increasing the fitness of others. If genes that promote altruism are shared by the beneficiaries of the altruism, these genes may increase in frequency. This mechanism could favor evolution of sacrifice for the benefit of relatives.

As one would expect, the necessity of considering both donors and recipients imposes a certain inherent complexity on models for the evolution of altruism. However, for the important case of parental altruism, it is possible to formulate a relatively simple model.

Bidding goodby to our zebras, let us consider an avian species where both parents make equally important contributions to nestlings' survival. Their activities in this regard (incubating eggs, feeding chicks, guarding the nest, etc.) endanger their personal survival and are thus altruistic. Suppose, for simplicity, that the entire complex of parental behavior is controlled by one locus with two alleles, A_1 and A_2 (our usual assumption). We now assume, however, that there are two parameters, u_{ii} and v_{ii} , associated with the genotype A_iA_i . The parameter u_{ii} , called viability, is the probability that an A_iA_i survives to reproductive age, given that he survives until he leaves the nest. The other parameter, v_{ij} , called nurturance, is a measure of the quality of parental care provided by A₁A₁. Specifically, $v_{ii}v_{km}$ is the probability that an offspring of A_iA_i and A_kA_m parents survives until he leaves the nest. The v parameters in models of this kind are usually described as fertilities [9], but it is understood that what I call nurturance is an important component of genetic fertility parameters [10, p. 51]. For my present illustrative purposes, I am essentially assuming that nurturance is the only component of fertility.

Let p_n be A_1 gene frequency among birds leaving the nest in the *n*th generation. It is not difficult to show that this quantity satisfies (2.1), if we take

$$w_{ii} = u_{ii}v_{ij} = \text{viability} \times \text{nurturance}.$$
 (4.1)

In this model, an altruistic individual is one with high nurturance and low viability, while a selfish individual has low nurturance and high viability. It is not too surprising that the course of evolution is controlled by the product of viability and nurturance, which I will call composite fitness.

194

To facilitate comparison with Hamilton's theory, let me rewrite (2.1) in the form

$$\Delta p_n = p_n(w_1 - w_{-})/w_{-}, \tag{4.2}$$

where $\Delta p_n = p_{n+1} - p_n$, $w_1 = w_{11}p_n + w_{12}q_n$, and $w_n = w_{11}p_n^2 + 2w_{12}p_nq_n + w_{22}q_n^2$. The theory in Hamilton's basic paper [11] covers relatives of all kinds, though he says that it is especially appropriate for interactions between relatives of the same generation. Using an argument involving "certain lapses from mathematical rigour" [11, p. 2], Hamilton obtains the equation

$$\Delta p_n = p_n(R_1 - R_1) / (R_1 + \delta S_1). \tag{4.3}$$

(This is Hamilton's equation (2) [11, p. 6], with i = 1 and dot superscripts omitted.)

Hamilton draws attention to the similarity between (4.3) and its classical analog, (4.2). Apart from the term $\delta S_{...}$, which we may ignore, the only difference between (4.2) and (4.3) is that, in place of composite fitness, w_{ij} , Hamilton has inclusive fitness, R_{ij} . This quantity is the sum of an individual's personal fitness, taking account of his altruistic or selfish acts, and the changes in relatives' fitnesses due to these acts. The latter changes are weighted by the appropriate coefficient of relationship, r, which is the expected proportion of genes of the donor and recipient that are identical by descent (that is, copies of the same gene in some recent ancestor). For parent and offspring, this coefficient is $\frac{1}{2}$, so

$$R_{ij} = 1 + x_{ij} + \frac{1}{2}y_{ij}, \tag{4.4}$$

where x_{ij} and y_{ij} are, respectively, the parent's and offsprings' changes in personal fitness due to the parent's altruistic or selfish behavior. In the absence of such behavior, we are assuming, for simplicity, that all genotypes have the same fitness, which we have taken to be 1 without loss of generality.

We now come to the main point of this section. According to Hamilton, "With classical selection a genotype may be regarded as positively selected if its fitness is above the average and as counter-selected if it is below." In the case of altruistic and selfish behavior, "the kind of selection may be considered determined by whether the inclusive fitness of a genotype is above or below average" [11, p. 14]. For easy reference below, I will call this Hamilton's criterion. We saw in §2 that it is incorrect in the classical case. Superior w_{ij} does not ensure that A_iA_j increases in frequency if $i \neq j$. The analogy between (4.2) and (4.3) leaves no doubt that Hamilton's criterion is also inconsistent with Hamilton's basic theory as expressed by (4.3). Perhaps this is what Hamilton had in mind when, in a later paper [12, p. 196], he described this criterion as "approximate."

Given the conflict between Hamilton's criterion and his basic theory, it is of considerable interest that his most frequently cited result is derived from the criterion. I am referring to the rule relating k to 1/r [11, p. 16], where k is the ratio of recipients' and donor's changes in fitness, $k = y_y/x_y$. In cases of

selfishness and altruism, x_{ij} and y_{ij} have different signs, so k < 0. For parent-offspring interaction, (4.4) yields

$$R_{ij} = 1 + (1 + k/2)x_{ij}. (4.5)$$

For other relationships, the comparable equation is

$$R_{ij}=1+(1+rk)x_{ij}.$$

Thus, according to Hamilton's criterion, natural selection should tend to favor an altruistic genotype $(x_{ij} < 0)$ if |k| > 1/r. Similarly, a selfish genotype $(x_{ij} > 0)$ should be favored if the reverse inequality holds.

These calculations with inclusive fitness, R, can be imitated with composite fitness, w. Recall that u_{ij} and $v_{ij}v_{km}$ are in units of offspring survival probability; hence v_{ij}^2 , not v_{ij} , is comparable to u_{ij} . This suggests rewriting (4.1) in the form $w = u(v^2)^{1/2}$ or

$$\log w = \log u + \frac{1}{2} \log v^2. \tag{4.6}$$

(Here and below we omit subscripts.) Only ratios of us, us, and ws are relevant to (4.2); hence these parameters can be rescaled so that a neutral condition (no selfishness of altruism) corresponds to u = 1, v = 1, and thus w = 1. Then the neutral values of the logarithms in (4.6) are 0, and these quantities are fully comparable to R = 1, x, and y in (4.4). Taking

$$k = \log v^2 / \log u, \tag{4.7}$$

we obtain

$$\log w = (1 + k/2)\log u, \tag{4.8}$$

the analog of (4.5). Altruistic (u < 1, v > 1) or selfish (u > 1, v < 1) genotypes satisfy Hamilton's criterion for positive selection (large w) depending on whether |k| > 2 or |k| < 2, just as in the inclusive fitness formulation.

To see that this rule need not correctly predict the course of evolution, we return once again to the example described in Table 1 and assume that A_1A_1 is neither altruistic nor selfish. Rescaling to achieve $w_{11} = 1$, we obtain $w_{22} = 1/3$ and $w_{12} = 4/3$. As a consequence of (4.7) and (4.8), $u = w^{2/(k+2)}$ and $v = w^{k/(k+2)}$. For k = -4, these formulas yield the values given in Table 2. According to

TABLE 2 Viability, nurturance, and asymptotic frequency for k = -4

Genotype	и	υ	w	Frequency
A ₂ A ₂ (selfish)	3	1/9	1/3	1/16
A_1A_1 (neutral)	1	1	1	9/16
A_1A_2 (altruistic)	3/4	16/9	4/3	6/16

Hamilton's rule, A_1A_2 altruists should prosper, since |k| > 2, but, in fact, selection favors the neutral genotype, A_1A_1 . Table 3 has been obtained by reversing entries in the u and v columns of Table 2. This transformation preserves w but changes k to k' = 4/k = -1.

196

Genotype	u	v ·	w	Frequency
A ₂ A ₂ (altruistic)	1/9	3	1/3	1/16
A_1A_1 (neutral)	1	1	1	9/16
A_1A_2 (selfish)	16/9	3/4	4/3	6/16

Since |k| < 2, the selfish genotype should be favored, but, again, it is asymptotically less prevalent than the neutral genotype. For further consideration of "|k| > 1/r," in the context of other models, see [13] and [14].

5. Genes and genotypes. Fitnesses of genes can be defined by averaging over genotypes in which they occur. In the classical case, the fitnesses of A_1 and A_2 are

$$w_1 = w_{11}p_n + w_{12}q_n$$
 and $w_2 = w_{12}p_n + w_{22}q_n$

Although (4.2) and (4.3) do not imply positive selection of genotypes with above average fitness, they do imply positive selection of genes with above average fitness. Thus discussions of evolution in terms of genes rather than genotypes avoid the pitfall which this paper has been concerned.

Many discussions of evolution (e.g., parts of [15]) are, in fact, cast in terms of genes. A limitation of this approach is that the fitness of a gene is not as intuitive as the fitness of a genotype. It is frequency dependent, and, if the heterozygote is most (or least) fit, even the ordering of w_1 , and w_2 is different for large and small p_n . Discussions that ignore this frequency dependence implicitly (and perhaps unconsciously) rule out heterozygote superiority, which is the basis for the example considered repeatedly in this paper.

References

- 1. M. F. Norman, A "psychological" proof that certain Markov semigroups preserve differentiability, these PROCEEDINGS.
 - 2. E. O. Wilson, Sociobiology: the new synthesis, Harvard Univ. Press, Cambridge, Mass., 1975.
 - 3. W. J. Ewens, Mathematical population genetics, Springer-Verlag, New York, 1979.
 - 4. S. J. Gould, Ever since Darwin: reflections in natural history, Norton, New York, 1977.
 - 5. J. Maynard Smith, Evolution and the theory of games, Amer. Sci. 64 (1976), 41-45.
- 6. G. F. Oster and S. M. Rocklin, Optimization models in ecolutionary biology, Some Mathematical Ouestions in Biology, vol. 10 (S.A. Levin, ed.), Amer. Math. Soc., Providence, R. I., 1979.
- 7. P. D. Taylor and L. B. Jonker, Evolutionarily stable strategies and game dynamics, Math. Biosci. 40 (1978), 145-156.
- 8. J. Hofbauer, P. Schuster and K. Sigmund, A note on evolutionary stable strategies and game dynamics, J. Theoret, Biol. 81 (1979), 609-612.
 - 9. W. F. Bodmer, Differential fertility in population genetics models, Genetics 51 (1965), 411-424.
 - 10. T. Nagylaki, Selection in one- and two-locus systems, Springer-Verlag, New York, 1977.

- 11. W. D. Hamilton, The genetical evolution of social behavior. I, J. Theoret. Biol. 7 (1964), 1-16.
- 12. _____, Altruism and related phenomena, mainly in social insects, Ann. Rev. Ecology and Systematics 3 (1972), 193-232.
- 13. E. L. Charnov, An elementary treatment of the genetical theory of kin-selection, J. Theoret. Biol. 66 (1977), 541-550.
- 14. L. L. Cavalli-Sforza and M. W. Feldman, Darwinian selection and "altruism," Theoret. Pop. Biol. 14 (1978), 268-280.
- 15. R. Dawkins, The selfish gene, Oxford, New York, 1976.

DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF PENNSYLVANIA, PHILADELPHIA, PENNSYLVANIA 19104