Components of fitness become effectively neutral in equilibrium populations

(neutralism/selectionism/soft selection/rank-order selection)

RUSSELL REEVE*, ERIC SMITH*, AND BRUCE WALLACE†

Departments of *Statistics and †Biology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

Contributed by Bruce Wallace, December 26, 1989

ABSTRACT If, because of trade-offs or similar considerations, components of fitness are negatively correlated in equilibrium populations, they may then exhibit no correlation with total fitness. Lack of correlation with fitness is a fundamental characteristic of selective neutrality. Hence, selectively important traits become subsumed under neutral traits when the definition of neutrality is extended to include factors (genetic or phenotypic) that "behave as if they are neutral."

The matter of neutrality with respect to natural selection can be addressed at both the phenotypic and genotypic levels. The former usually involves intuitive judgments. Dobzhansky (ref. 1, p. 261), for example, asks why the anterior scutellar bristles in many species of Drosophila are convergent but in others are divergent. Lewontin (see ref. 2, p. 895) cites one- and two-horned rhinoceroses as exhibiting phenotypes whose differences were established by chance—i.e., the attributes one versus two horns are neutral with respect to natural selection.

At the genetic level, the neutralism of alternative alleles implies that they lie below the stratum at which natural selection is effective and, hence, are interchangeable. A rigorous definition of neutrality at this level might read: two alleles are selectively neutral if either could be substituted for the other in any cell, tissue, or organ without affecting the altered individual's prospect of survival and reproduction under any environmental circumstance. Such a rigorous definition has not been adopted. Rather, neutrality has been extended to include instances (i) in which selection intensity affecting the alternative alleles fluctuates and (ii) in which chance events predominate in determining the fate of an allele in a population (ref. 3, p. 165). Thus, alleles with either small advantages or disadvantages in a large population would become effectively neutral in a small one.

The present essay attempts to demonstrate that under the extended view of neutrality, even important components of fitness are effectively neutral in equilibrium populations. That the genetic bases of neutral phenotypes (hidden as they are beneath these phenotypes) are also neutral is assumed without formal proof.

Fitness is an extremely complex trait that involves survival and reproduction, each of which hinges, in turn, on a multitude of smaller components. Survival among animals, for example, may depend on the avoidance of predation. This, in turn, may be affected by speed, agility, camouflage, or patterns of activity. Again, each of the factors affecting survival is a complex of many subcomponents. Speed, for example, is sufficiently complex that breeders of race horses in Britain have been unable to lower steeply chases over the past 50 years (4). Camouflage, depending on it does more on impressionism than realism, is also an extremely complex matter. And so it is with seasonal activity; enumerating all the variable factors involved in an individual's behavior would be an impossible task.

Any aspect of the phenotypes that is determined by the interaction of many components must, at best, be poorly correlated with any one component. As a purely intellectual exercise, consider 10 columns of randomly chosen digits (0-9) with 20 digits constituting each column. Consider, too, the sums of the 10 digits in each row. What is the expected correlation between the column of sums and the digits that constitute any one column? Clearly, the correlation cannot be perfect. In fact, the expected correlation, r, equals (1/n)½ where n equals the number of columns (see Appendix). In the cited example involving 10 columns, the expected value of r would be 0.32. If dozens of columns were to be summed, however, the correlation between the sums of rows and the items constituting any single column would become extremely small [(1/n)½ would approach zero] even though no column is more important than any other in determining the sums.

Consider, now, a somewhat different example: Rather than consisting of random digits, the individual columns are now themselves correlated with correlation γ. The expected correlation (see Appendix), of the sums with any one column, under this stipulation,

\[ r = \sqrt{(1-\gamma)/n + \gamma}. \]  

[1]

Consequently, if γ equals 1.00, r equals 1.00; if γ = 0, r equals (1/n)½, as before.

Of more interest, however, is the condition under which r equals 0. This occurs when the digits in the individual columns are negatively correlated:

\[ \gamma = -\frac{1}{n-1}. \]  

[2]

The negative correlation, in this case, does not involve the square root of the number of columns but, essentially, the number itself. By analogy, fitness will not be correlated with any of its individual components if the latter are, overall, negatively correlated with one another. Any aspect of the phenotype, however, that is not correlated with fitness, is, by definition, neutral.

That components of fitness may be negatively correlated was emphasized 30 years ago by Robertson (5). If one can reveal within a population the presence of additive genetic variation for any component of fitness, one can draw either of two conclusions: (i) The population is not at equilibrium. (ii) If it is at equilibrium, then components of fitness must be negatively correlated. Examples illustrating the second possibility are so plentiful that the first is virtually ignored, even though one must admit that populations are at times displaced from equilibrium conditions. The following can be cited as examples of negative correlations that exist among components of fitness: Egg size versus egg number; longevity versus...
reproductive effort; percentage egg hatch versus speed of embryonic development (haste makes waste even at the level of molecules; see ref. 6); foraging time versus time spent protecting offspring; the need for concealment versus the need for sexual displays or the need to obtain food and water. The sheer number of negatively correlated components of fitness virtually assures that even though some components may be positively correlated (size of an herbivore and its avoidance of predation), overall fitness will not be correlated with any of its many components. All phenotypic traits, consequently, are neutral, even if they are not.

What bearing has the above proof on the neutralist-selectionist "controversy"? First, the goodness of fit between predictions made under neutralist assumptions and conditions prevailing in natural populations may suggest that (at least before human intervention occurred on a grand scale) most populations of breeding adults have been at or near equilibrium conditions most of the time. Second, when neutralism was extended to include not only neutral alleles as defined rigorously but also alleles that "behave as if they are neutral," neutralism effectively included all phenotypic traits and all alleles that are responsible for the development of these traits.

The transformation of a component of fitness into an effectively neutral phenotypic trait is illustrated in Fig. 1. The horizontal axis of the diagram represents a component of fitness that is more-or-less normally distributed in a population (dashed line). Over much of its range, this component of fitness is directly correlated with fitness itself (solid line). However, at some point (because of negative correlations with other components of fitness and because of various trade-offs), this component exhibits a maximum fitness beyond which it declines. In short, there are intermediate optima even for components of fitness.

Selection in large measure is rank order. In Fig. 1, such selection would be accomplished by removing the lower portions of the fitness distribution, gradually ascending upward until the number of surviving individuals corresponds to that which the environment can "carry." By definition, the surviving individuals will exhibit maximum fitness. By projecting the limits of this range of (total) fitness onto the horizontal axis, one determines the population of immature individuals from which the surviving adults are selected (stippled area). It is within this population of adults that the lack of correlation between fitness and its components applies.

**APPENDIX**

In this appendix, we prove the results in the paper. Assume there are $n$ traits and let $X_1, X_2, ..., X_n$ be the measurements on the $n$ traits. Let the correlation between $X_i$ and $X_j$ be $\gamma_{ij}$. Let the average correlation between $X_1$ and $\{X_2, ..., X_n\}$ be $\gamma_1$; e.g.,

$$\bar{\gamma}_1 = \frac{\sum_{j=2}^{n} \gamma_{1j}}{n-1}.$$ 

And let $\bar{\gamma}$ be the average of all correlations; e.g.,

$$\bar{\gamma} = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \gamma_{ij}}{n(n-1)/2}.$$ 

Then (assuming equal variances)

$$\text{Var} \left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \sigma^2 + 2\sigma^2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \gamma_{ij}$$

$$= n\sigma^2 + \sigma^2 n(n-1)\bar{\gamma}.$$ 

Also,

$$\text{Cov} \left( X_1, \sum_{i=1}^{n} X_i \right) = \sigma^2 \sum_{i=2}^{n} \gamma_{1i}$$

$$= \sigma^2 [1 + (n-1)\bar{\gamma}_1].$$

So,

$$\rho \left( X_1, \sum_{i=1}^{n} X_i \right) = \frac{\sigma^2[1 + (n-1)\bar{\gamma}_1]}{\sigma^2[n(1 + (n-1)\bar{\gamma})]^{1/2}}$$

$$= \frac{1 + (n-1)\bar{\gamma}_1}{\sqrt{n} \sqrt{1 + (n-1)\bar{\gamma}}}.$$ 

Suppose that $X_1$ acts neutral, so $\rho(X_1, \sum_{i=1}^{n} X_i) = 0$. We then get

$$\bar{\gamma}_1 = -\frac{1}{n-1}$$

or

$$\sum_{j=2}^{n} \gamma_{1j} = -1.$$ 

Hence, the requirement is that the sum of the correlations between $X_1$ and the other variables is $-1$. One may also

---

*Fig. 1. Rank-order selection with respect to a component of fitness that possesses an intermediate (albeit displaced toward the high end) optimum. Dashed line illustrates the frequency distribution of the component of fitness among individual members of the population. Solid line illustrates (i) the correlation between fitness and the component of fitness over much of the latter's range, (ii) the optimum value of the component relative to total fitness, and (iii) the negative impact on total fitness characterizing a component that is overdeveloped at the expense of other, equally important components. Stippled area represents the zygotes (equal in number to the carrying capacity of the environment) whose fitnesses exceed that required for survival.*
interpret this as the sum of the covariances is the negative of
the variance. Also note that the result does not depend on the
correlations not involving $X_1$.

In the equicorrelated case, we have ($\gamma_0 = \gamma$ for all $i, j$)

\[
\rho(X_1, \sum_{i=1}^{n} X_i) = \frac{\sqrt{1 + (n - 1)\gamma}}{\sqrt{n}}
\]

\[
= \sqrt{\frac{1 + n - 1}{n}} \gamma
\]

\[
= \sqrt{\frac{1 - \gamma}{n} + \gamma},
\]

which is Eq. 1. Setting $\rho$ to 0, we obtain Eq. 2,

\[
\gamma = -\frac{1}{n - 1}.
\]

Notice that if $\gamma_0 = 0$ for all $i, j$, the correlations above reduce
to

\[
\rho(X_1, \sum_{i=1}^{n} X_i) = \frac{1}{\sqrt{n}},
\]

which is well known. The above result can be readily ex-
tended to situations with unequal variance or the situation
where the traits are weighted.

Valuable comments by R. C. Lewontin and Rollin Richmond are
gratefully acknowledged.

1. Dobzhansky, Th. (1970) Genetics of the Evolutionary Process
(Columbia Univ. Press, New York).
Introduction to Genetic Analysis (Freeman, San Francisco).
Univ. Press, New York).
20, 225–229.