

# HARD AND SOFT SELECTION REVISITED<sup>1</sup>

BRUCE WALLACE

*Department of Genetics, Development, and Physiology, Cornell University, Ithaca, N. Y. 14850*

Received November 27, 1974

Revised June 17, 1975

In June 1967 (see Wallace, 1968) at a symposium sponsored by Syracuse University and the New York State Science and Technology Foundation, I first used the terms "hard" and "soft" selection. Since then these terms have enjoyed a modest popularity. Not always, however, have they been used in the sense I originally intended. Such misuse is not unexpected because the terms, although illustrated by examples, were not defined. Admittedly, my own interpretation of these terms has changed during the intervening years. It seems appropriate, therefore, to review the reasons for which the terms *hard* and *soft* were first used, to identify instances in which they have been obviously misapplied or misunderstood, and, using the original examples, to develop definitions for them. Finally, I want to examine Clarke's (1973*a, b*) alpha-selection to see whether, as he suggests, it corresponds to my soft selection.

## THE ORIGINS OF THE TERMS: *HARD* AND *SOFT* SELECTION

The terms *hard* and *soft* have origins that trace to the concept of genetic load. Genetic load, it may be recalled, is defined (Crow, 1958) as the proportional amount by which the average fitness (or any other measurable trait) of a population is reduced *relative to that of the optimal genotype*. The genetic load concept as defined leads almost inescapably to the conclusion that populations of small or medium size are genetically homogeneous (Kimura and Crow, 1964): many neutral

alleles cannot be maintained in a population consisting of hundreds or thousands of individuals; alternatively, the selective elimination of ill-adapted homozygotes supposedly imposes an unbearable load on a population if alleles at more than a few loci are retained through the selective superiority of heterozygous individuals.

That populations are not genetically homogeneous but, on the contrary, are polymorphic for two or more alleles at one third, one half, or even more of all gene loci was revealed by studies utilizing starch-gel electrophoresis (Harris, 1966; Lewontin and Hubby, 1966). These observations stood in stark contrast to expectations based on genetic load theory. To accommodate them, one school (whose members are now known as the neutralists) merely increased the effective population size of natural populations; populations consisting of millions of individuals can retain selectively neutral alleles at high frequencies at many loci through mutation pressure alone. A second school, whose members are known as selectionists, holds that the observed genetic variability is maintained by selection (largely, but not necessarily, the superiority of heterozygous individuals). These persons, of necessity, reject one or more assumptions upon which genetic load calculations are based (see Sved et al., 1967; King, 1967; Milkman, 1967).

To avoid the intolerable loss of zygotes arrived at by the usual genetic (=segregational) load calculations, various devices have been sought. Sved and his colleagues suggested that for intrinsic physiological reasons an upper limit might be assigned to an individual's fitness. Milkman stressed

<sup>1</sup> This paper was prepared while the author's research was supported under contract No. AT-(11-1)-3149, U.S. Atomic Energy Commission.

that the individual, not the gene locus, is the unit of selection and, like King, argued that selection culls individuals from the low end of the fitness distribution, thus leaving as survivors those in the upper tail. The device that I developed under the designation soft selection seems to resemble most closely the ideas of Milkman and King.

The words *hard* and *soft* were borrowed not from population biology but from international monetary exchange. Certain countries operate on soft currency. When exchanging such currency for dollars (until recently, at least) it appears to be virtually worthless. Within the home country, however, financial transactions proceed with dispatch, goods are bought and sold, and the softness of the currency defies detection.

If the suboptimal fitness of certain genotypes depends upon the actual presence of individuals possessing the optimal genotype (as defined by Crow), the designation "suboptimal" in respect to fitness resembles "soft" in respect to currency. In the absence of the optimal genotype, those with the highest fitnesses (rather than an abstraction) become the real standards of comparison; individuals with these genotypes tend to survive while other, less fit, individuals are eliminated from the population. The inferior fitness of certain genotypes is revealed by a direct comparison with others possessing higher fitnesses, not by a mathematician's calculation.

In contrast to the above, lethal genes often kill their carriers under all known conditions. To calculate the proportion of individuals that will survive the combined effects of two, three, four, or more lethal genes, one need only obtain the product of the individual probabilities. The outcome is entirely reliable; it can be verified experimentally with relative ease. This type of selection I have called hard. The early calculations which were made in the name of genetic load were made according to rules which I regard as appropriate only in the case of hard selection.

#### DENSITY-DEPENDENT AND DENSITY-INDEPENDENT SELECTION

Because soft selection leads to a probability of survival and reproduction for one genotype that depends upon the presence or absence of individuals of other genotypes and because hard selection leads to a probability of survival and reproduction that is independent of these other individuals, the terms soft and hard selection resemble in at least some respects the older terms used by ecologists, density-dependent and density-independent selection. I might have used these older terms except that I was told they were controversial. Because I was trying to resolve what at that time had become the genetic load controversy, it seemed pointless to become unnecessarily entangled in an additional one. This consideration provided further reason for using the non-controversial, undefined terms *hard* and *soft*.

In retrospect, the controversy among some ecologists over the meanings of density-dependent and density-independent selection (if it still exists) seems resolvable. An example illustrating the nature of the conflict as I understand it goes as follows. The death of individuals who fail to establish private territories is density-dependent; had fewer individuals existed, all individuals could have had territories. In contrast, death caused by a capricious event such as a torrential rain is density-independent because such an event is independent of the numbers of individuals of earth-bound species. On second thought, however, suppose the rain kills only those individuals who have no territories. Does that make the deaths by rain density-dependent? (This example could have been inverted because the number of territories might depend upon whether or not a torrential rain has occurred.)

If the above example illustrates the controversy fairly, there really is none. A torrential rain, an early frost, or any other natural calamity that is unrelated to the number (or density) of individuals of a species may be *necessary* for density-

independent selection but is not *sufficient* in itself to guarantee that the resulting mortality is in fact density-independent. It is not appropriate to ascribe causes of death to individual factors which operate sequentially or synergistically.

Once again we can ask: "Should the terms hard and soft selection, be replaced by the older ones, density-independent and density-dependent selection?" Before addressing this question, I want to dispose of some obvious misuses of the undefined terms, misuses that clearly ignore the original illustrative examples.

#### MISUSE OF THE TERMS "HARD" AND "SOFT" SELECTION

In at least one publication, egg and larval mortality have been equated with soft selection. This is a misuse of the term. Lethal genes, for example, provided an illustration of hard selection: one balanced-lethal system kills one half of all individuals and permits one half to survive; two independent balanced-lethal systems permit only one quarter of all zygotes to survive; three, one eighth; four, one sixteenth; and so forth. This example illustrates the mathematical rigor that characterizes hard selection; furthermore, it serves as an example of hard selection even though the lethal genes involved might kill individuals very early in development. Juvenile deaths may raise questions about the replacement of missing individuals (see Wallace, 1963) but in themselves do not serve to distinguish between hard and soft selection.

Dobzhansky (1970: 226-277) has suggested that the terms *hard* and *soft* be replaced by the terms *rigid* and *flexible*. The latter terms have been used (Dobzhansky, 1962) in describing the contrasting patterns of chromosomal polymorphisms in *Drosophila*. Although the two pairs of terms have a superficial similarity, they describe quite different things and, therefore, are not interchangeable.

The terms "rigid" and "flexible" were first used to describe the gene frequencies

(or, in Dobzhansky's studies, the frequencies of cytologically identifiable gene arrangements) observed in samples obtained from spatially isolated populations or from a single population that was sampled periodically over a prolonged time. If the arrays are statistically homogeneous, the polymorphism is said to be "rigid"; if, on the other hand, the arrays are statistically heterogeneous, the polymorphism is said to be "flexible." These terms have been used in precisely this original sense by Crumpacker and Williams (1974) in reference to populations of *D. pseudoobscura* found in Colorado. For unknown reasons, inversion frequencies in populations living south of Denver undergo seasonal cycles (i.e., are flexible) whereas those north of the city do not (i.e., are rigid). It is difficult, if not impossible, to equate the terms *rigid* and *flexible* which describe the alternative outcomes of a statistical test for homogeneity with *hard* and *soft* which refer to the action of selection within a population.

#### ARE HARD AND SOFT SELECTION IDENTICAL TO DENSITY-INDEPENDENT AND DENSITY-DEPENDENT SELECTION?

The non-committal (and undefined) terms, hard and soft, were intended to cover some of the situations which are known to ecologists and ecological geneticists as density-independent and density-dependent; otherwise, I would not have considered using the older terminology. Nevertheless, the existence of the two terminologies provides an opportunity to compare them and, should they prove not to be identical, to retain both while making clear what their differences are. In anticipating the outcome of this comparison, I expect to demonstrate that soft selection differs from density-dependent selection by including *frequency*-dependent selection as well. That is, I expect to show that my earlier examples of soft selection (Wallace, 1968, 1970) depend upon both density of individuals and the frequency of differing genotypes unlike density-dependent selec-

tion which may be independent of gene frequencies.

Hard selection is that which can be predicted (as population geneticists have always done) with no recourse to individuals. The selective coefficient for a particular genotype is fixed. This coefficient holds for all gene frequencies; therefore, one can perform calculations that predict the elimination of a deleterious gene from a population or the establishment of an equilibrium between mutation and selection. Density, in the sense of numbers of individuals per unit space, never enters into these calculations. Consequently, hard selection is both density- and frequency-independent. For that reason, it is the mathematical geneticist's dream; for the same reason, it is the ecologist's nightmare. Calculations can be carried out without qualification over all gene frequencies from 1.00 to 0 and, through the use of either additive or multiplicative processes, over any number (from 1 to 10,000 or more) of gene loci. Unfortunately, little of the real world corresponds to calculations of this sort.

I suggested above that soft selection is both density- and frequency-dependent. This suggestion needs closer scrutiny using as evidence the examples which I cited in the past. If the suggestion is borne out, I shall propose that density-dependent, frequency-dependent selection be referred to as soft selection.

One example (Wallace, 1968, 1970: 90) involved the seemingly fixed number of suitable pupation sites in some *Drosophila* culture vials; these vials tend to yield constant numbers of adults. Details might vary somewhat in respect to the survival of particular genotypes out of the arrays generated by different laboratory strains but the implications are quite clear: any strain, or population, of *Drosophila* flies grown in these vials would yield essentially the same number of adults because a fixed number of suitable pupation sites is available to them. I would expect, for example, that a wild-type strain of *D. melanogaster*

and a mutant strain carrying several marker genes to yield about the same number of flies in such cultures. And, in a sense, it is known that such different strains do yield remarkably similar numbers of flies. The relative ease with which homozygous mutant flies can be maintained in stock cultures fails to reflect the low survival values these same genotypes would exhibit were they to develop together with wild-type flies in a single culture bottle. In many respects an adult fly resembles the Dean of the Faculty at a university: the position exists and needs filling; the creation of a dean does not require, however, the existence of faculty members with deanlike qualities because the position will be filled in any case.

The example cited above has obvious bearing on density-dependent selection. The number of survivors is more-or-less fixed by suitable pupation sites. Consequently, one can imagine cultures in which the probability of survival approaches 100% (when the number of young larvae is equal to or less than the number of pupation sites) or exceedingly low (when the number of young larvae greatly exceeds the number of these sites). Now, if among the competing larvae, genotypes of varying fitnesses (with respect to successful pupation) occur in *constant* proportions, then the greater the number of young larvae in the culture, the greater the probability that all survivors will be of the more successful genotypes. The relative frequencies of different genotypes in this instance were said to be constant; therefore, the poorer chance of survival of the weaker genotypes and the greater proportion of the stronger ones among adult survivors represent changes wrought by *density*. Hence, these cultures (and all other situations in which the number of survivors is largely determined by sites or territories) illustrate the working of density-dependent selection.

Consider now, however, the same culture conditions as before but in the present case the number of young larvae competing for pupation sites is held constant—say, four

times as many as can successfully pupate. In this example, we now imagine that there is a greatly superior genotype which (in different cultures) is represented by 0, 1, 2, 3, 4, . . . individuals. Clearly, the survival of larvae of the less successful genotypes will depend upon the frequency of superior larvae. Density, in this case, has been held constant while the relative frequencies of different genotypes were allowed to vary. Variation in frequency affects survival; consequently, soft-selection is frequency-dependent selection. Thus, by the arguments presented in these two paragraphs, the example cited as an illustration of soft selection has been shown to involve both density- and frequency-dependent selection.

A second example used in illustrating soft selection (Wallace, 1970: 91-93) involved not survival but the reproductive (mating) success of male flies. By means of a simple experiment whose details need not be repeated here, male *D. melanogaster* raised in laboratory population cages were shown to be only 1/16th as efficient at inseminating females as were bottle-raised males. "Cage-raised" and "bottle-raised" in this illustration can be regarded as flies of two contrasting genotypes. Had the study involved relative viability, 1/16 would have qualified for the designation "lethal", thus, the cage-raised males would have been treated as dead. Obviously, however, these males keep cage populations going. In the absence of the more efficient (bottle-raised) males, there is no detriment in the tardiness with which cage males find and inseminate females. The experimentally demonstrated "inadequacy" of these males in no way threatens the existence of cage populations.

The above example was used to illustrate soft selection; the two types of males can as well be thought of as two contrasting genotypes. The designation "soft selection" was based in this instance on frequency-dependent selection: the presence of superior males revealed an inadequacy not previously suspected; their presence in

sufficient proportions might have prevented the cage-reared males from fathering any offspring at all.

Suppose that one did not alter the relative proportions of the two types of males but, rather, altered the total numbers of each while keeping their proportions constant. Would the increase in total numbers of males decrease the probability of mating on the part of the inferior males? Not if the total number of females is correspondingly increased! Thus, (regarding "cage-raised" and "bottle-raised" as designations for two genotypes), 500 females + 500 cage-raised males + 100 bottle-raised males should confer the same relative fitness on cage males as would 1000 females + 1000 cage-raised males + 200 bottle-raised males. If density depends upon population number and if the numbers of the two sexes fluctuate up and down together, the differential mating ability of males may illustrate frequency-dependent (density-independent) selection. In this case, the example does *not* illustrate soft-selection as I now define this term. On the other hand, if the number of females were artificially held at 500 while the males were increased to 1000 (cage) + 200 (bottle) from an earlier 500 (cage) + 100 (bottle), mating success would prove to be density-dependent as well as frequency dependent. Females in the latter case would correspond to the pupation sites or territories of the earlier example.

A third example of the ostensible working of soft selection is described in Wallace (1970: 106) and has been studied experimentally by Coman and Wallace (1973). The original claim was as follows: if members of each generation serve as their own competitors and as their own standards, inbreeding might *appear* to lead to a temporary depression in fitness which would later *appear* to be reversed. The explanation for this hypothetical course of events is that the variance in fitness among individuals might increase at first and then decrease. Both at the start of the inbreeding program (where the starting

population consists of uniform  $F_1$  hybrid individuals) and late in its course (when homozygosity would have been established by inbreeding), the between-individual variance would be small and the expected selective mortality would also be small. Small mortality can, of course, be misinterpreted as high survival or high fitness. In the intervening generations, the between-individual variance would be larger, selective mortality would be larger, over-all survival in the culture lower, and the (erroneous) estimate of fitness lower.

The data presented by Coman and Wallace (1973) closely followed the predicted pattern; however, these authors were reluctant to claim that the suggested interpretation of this course of events is correct. The present task, however, is to re-examine the original argument in the light of the definitions of hard and soft selection that are now being developed; that is, that soft selection is both density and frequency dependent whereas hard selection is neither.

The argument concerning the role of variation in fitness and selective mortality can be best visualized in terms of territories even though the immediate example involves *Drosophila* larvae in a culture vial. Neglecting events that might occur under conditions of pathological over-crowding (where all individuals might starve), larvae of uniform food gathering abilities in a crowded culture bottle can be imagined as eking out livelihoods in rather minimal, three-dimensional territories that permit the greatest number of larvae to survive. This expectation would hold whether the larvae are uniformly vigorous or uniformly decrepit.

If, in contrast to the preceding stipulation, the population of larvae is highly variable in respect to the vigor with which they seek and compete for food, some larvae may obtain more than the minimal amount of food required for surviving, thus assuring their own excellent health as adults. Weaker larvae, by the same token, may be prevented from obtaining enough food to

		DENSITY	
		INDEPENDENT	DEPENDENT
FREQUENCY	INDEPENDENT	HARD SELECTION	DENSITY DEPENDENT SELECTION
	DEPENDENT	FREQUENCY DEPENDENT SELECTION	SOFT SELECTION

FIG. 1. A diagram illustrating the relationship between four types of selection: hard, soft, density-dependent, and frequency-dependent.

survive. Consequently, variation among larvae in competitive food-gathering ability is expected (in the absence of cross-feeding, facilitation, or other complications) to reduce the number of surviving adults.

Whether in fact the survival of larvae follows the scheme outlined above or not and, furthermore, whether this scheme explains the results obtained by Coman and Wallace (1973), the example as described fits the new definition of soft selection. The postulated elimination of weak and poorly competitive larvae would be dependent upon both the frequency of aggressive ones among a larval population of constant size, and the total density of the larval population with the frequency of aggressive ones held constant.

This brief review of the examples which were cited earlier (in lieu of a formal definition) to illustrate soft selection has shown that the terms hard and soft selection have meanings that complement the older density-dependent (or independent) and frequency-dependent (or independent) concepts in an intellectually pleasing manner. As shown in Figure 1, hard selection is the absence of density and frequency dependence whereas soft selection is the





