SELECTION IN REFERENCE TO BIOLOGICAL GROUPS

I. INDIVIDUAL AND GROUP SELECTION APPLIED TO POPULATIONS OF UNORDERED GROUPS

By B. GRIFFING*

[*Manuscript received August* 18, 1966]

Summary

A selection theory designed to accommodate interactions among genotypes is presented. This involves defining unordered groups of genotypes among which interactions may occur, and describing populations of groups generated as combinatorial products of base populations. Gene models are developed which consist not only of direct contributions of the genotypes they represent but also of associate effects from other genotypes in the group.

With the use of group theory procedures, consequences of individual and group selection are investigated. It is found that individual selection is unbalanced in the sense that positive selection can result in a negative response of the population mean. On the other hand, it is demonstrated that group selection is balanced and as such invariably leads to desirable results. Thus positive group selection always produces non-negative responses, and continued selection results in a realization of the maximum potential of the population.

I. INTRODUCTION

Selection theory deals with changes in population structure due to differential reproduction of genotypes that constitute the population. Past developments in the theory have dealt with the consequences arising from extension of the gene model to accommodate almost any known transmissional genetic phenomenon. However, these developments have taken place within the framework of a single population in which genotypes do *not* interact.

In the present series of studies an attempt is made to take into consideration possible interactions among genotypes. If interactions are to be considered, the conceptual biological model used in development of the selection theory must be extended in two ways. First, groups of interacting genotypes must be defined and the population of such groups described. Second, the usual gene model for a given genotype must be extended to include not only *direct* effects of its own genes, but also *associate* contributions from other genotypes in the group. The objective, then, is to determine the consequences of selection operating on this more complicated, but at the same time more realistic, biological model of the real world.

Since this is the first paper in a series dealing with selection as it pertains to groups of interacting individuals, it is worth while to outline some of the problems that can be attacked by group theory approach. These range from the simplest problems having to do with competition within a *single* species to much more complex problems of interspecific dynamics when two or more species are involved.

* Department of Zoology and Entomology, Ohio State University, Columbus, Ohio.

As an example of one of the simplest, but as yet unsolved, problems involving a single species, consider the practical situation of breeding for increased yield of a densely planted crop. The greater the planting density the greater is the competition among plants for limited environmental space (nutrients, water, and light). A common manifestation of such competition is that the same genotype has different expressions in populations having different genotypic structures, e.g. in populations of mixed genotypes as contrasted with populations of pure stands. Obviously this leads to a plant-breeding dilemma since selection necessarily operates on a *mixed* population but has as a goal the production of the highest yielding *pure stand.*

An elegant analysis of this type of problem was made by Wiebe, Petr, and Stevens (1963) in their study of interactions among certain barley genotypes. Results of this analysis caused the authors to draw the following disturbing conclusions:

> "Significant reversals in relative yield were found to exist in comparisons between the same genotype, *VV* or *vv*, when grown in pure stand and in an advanced generation, thus indicating that the poorest plants should be saved from an advanced hybrid population rather than the good ones when yield is the criterion for selection. If this phenomenon has a degree of universality, then it may explain why breeding for increased yield has progressed so slowly."

The problem of responses varying in accordance with the genetic milieu is of immense importance, particularly in the breeding of small-grained cereals. This difficulty has undoubtedly been the reason why past selection results have been unpredictable and why breeding of small-grained crops has had the reputation of being an "art" rather than a scientific procedure for which definite rules could be formulated.

Another plant breeding problem at a higher level of complexity, but one which still involves only one basic species, is that of breeding for high yields of simple and complex mixtures. Studies of Suneson (1956) and Allard (1961) have caused a resurgence of interest in this form of breeding, particularly for self-fertilized crops. It has been repeatedly shown that mixtures exhibit greater phenotypic stability than components taken separately [see Simmonds (1962) for review]. It has also been suggested that use of mixtures may provide the answer to long-term adaptability to disease problems. However, for economical utilization of mixtures, the main task is to develop selection procedures which will integrate component parts of the mixture in such a way as to maximize the potential of the environmental space. This is a problem for which previously developed selection theory is inadequate. It requires a new approach based on interacting groups of genotypes.

Still considering a single species, the possibility exists of studying interactions among group elements which are not associated at random. An important example of this class of problems is that in which the group members are relatives. Group theory, then, provides the basis for a broad attack on problems due to interactions among relatives. This is a subject of considerable interest to animal breeders. As Hamilton (1964) has pointed out, a consideration of the interactions of relatives is also important in understanding the evolution of social behaviour.

Turning very briefly, now, to problems which encompass more than one species, it is hoped that the group theory approach will be of some assistance in analysis of the dynamics of a closely integrated system of interacting species, whether or not interactions are of a competitive or symbiotic nature. An example concerning artificial selection primarily is that of pasture breeding in which it is necessary to develop simple and complex mixtures of different species. In this case the basic problem is one of utilizing the symbiotic potential inherent in grass–legume combinations.

A more esoteric study, involving natural selection only, is that devoted to elucidation of processes which contribute to the ascent, decline, and final extinction of a species. This is a problem which must be studied within a framework of an interacting system of species. According to Lewontin (1965) it is one which has not received sufficient attention.

Finally, an interesting example involving both natural and artificial selection is that of the sometimes spectacular evolution of weeds in association with the development of crop plants. This evolution results as an interplay of natural and artificial selection operating on components of the crop–weed complex.

In all of these areas the group approach should prove useful in representing complicated biological situations more realistically, that is, in extending consideration from that of non-interacting individuals to that of groups containing individuals which may interact. By so doing, it is hoped to clarify some of the general problems of selection theory. In the more specific area of plant and animal breeding, the aim will be to identify selection procedures which will ensure that genetic change is in the most desirable direction and that continued selection will culminate in production of a population with the highest possible potential.

II. Consequences of Selection with regard to Unordered Groups

The subject matter of this first paper in the series concerns consequences of selection applied to the simplest kind of group, e.g. the unordered group whose members are randomly associated. The term "unordered" is used to indicate that different spatial configurations of individuals within a group in no way alter their genotypic values.

The following sections are concerned with construction of conceptual populations of groups varying in size from one and two members to an arbitrary number of members. Parameters in terms of a gene model and its associated variances are defined and consequences of various forms of selection are considered.

The theoretical analyses deal with an arbitrary number of alleles at a single autosomal locus. However, with groups of n individuals a complete description of the genetic situation generated by this single locus as it occurs among all members of the group requires an n-locus model. With regard to this model a separate arbitrary system of dominance parameters is assumed for direct and associate effects. Also separate arbitrary systems of epistatic effects are assumed for direct–associate and for associate–associate combinations of genotypes. With such a model the interesting situation occurs that the same allele may interact with itself as expressed in different individuals!

(a) Groups of Size One

Analysis of groups of size one are included in order to facilitate comparisons of past selection theory which ignores interactions among genotypes with that to be developed.

(i) *Population Parameters*

The conceptual population of groups of size one is identical to the genotypic array for the base population, i.e.

$$\sum_{i,j} p_i p_j (A_i A_j).$$

Let the genotypic value of A_iA_j, measured as a deviation from the population mean, be d_{ij}. This genotypic value is represented by the usual gene model

$$d_{ij} = a_i + a_j + \delta_{ij},$$

where

$$a_i = \sum_j p_j d_{ij} = \text{additive effect of the allele } A_i,$$

and

$$\delta_{ij} = d_{ij} - a_i - a_j = \text{dominance effect which is defined to be the interaction between alleles } A_i \text{ and } A_j.$$

The total genotypic variance may be partitioned as

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2,$$

where

$$\sigma_G^2 = \sum_{i,j} p_i p_j (d_{ij})^2 = \text{total genotypic variance},$$

$$\sigma_A^2 = 2 \sum_i p_i (a_i)^2 = \text{additive variance, and}$$

$$\sigma_D^2 = \sum_{i,j} p_i p_j (\delta_{ij})^2 = \text{dominance variance.}$$

(ii) *Consequences of Individual Selection*

Consequences of individual selection in a population of non-interacting genotypes are well known. Hence only pertinent results are listed.

(1) The selective value of A_iA_j is

$$w_{ij} = 1 + (\bar{i}/\sigma)_{\text{ind.}} d_{ij},$$

where $\bar{i}$ = standardized selection differential which is assumed to have a positive value, and σ = phenotypic standard deviation, and the subscript "ind." is used to indicate that $\bar{i}$ and σ relate to individual observations.

(2) The increment change in the frequency of A_i is

$$\Delta p_i = (\bar{i}/\sigma)_{\text{ind.}} a_i.$$

(3) The increment change in the population mean due to one cycle of selection is

$$\Delta \mu = (\bar{i}/\sigma)_{\text{ind.}} \sigma_A^2.$$

Since $\Delta\mu$ is a function of quantities which are all non-negative, selection continuously *increases* the population mean until the maximum value is attained.

(*b*) *Groups of Size Two*

(i) *Population Parameters*

Groups consisting of two genotypes are represented as the couplet $(A_{i_1}A_{j_1}, A_{i_2}A_{j_2})$. A population of such groups can be obtained as the two-way combinatorial product involving the genotypic array of the base population, i.e.

$$[\sum_{i,j} p_i p_j (A_i A_j)] \times [\sum_{i,j} p_i p_j (A_i A_j)] = \sum p_{i_1} p_{j_1} p_{i_2} p_{j_2} (A_{i_1}A_{j_1}, A_{i_2}A_{j_2}).$$

Because of the unordered property, groups having the same elements, regardless of order, are identical. Thus only one array of genotypic values needs to be considered.

The genotypic value of $A_{i_1}A_{j_1}$ as expressed in the group $(A_{i_1}A_{j_1}, A_{i_2}A_{j_2})$ is denoted as ${}_{i_1j_1}d_{i_2j_2}$ and coded so that

$$\sum p_{i_1} p_{j_1} p_{i_2} p_{j_2} ({}_{i_1j_1}d_{i_2j_2}) = 0.$$

Subscripts in front of the genotypic symbol '*d*' indicate the genetic constitution of the individual itself, whereas subscripts following '*d*' indicate the genetic constitution of the associated genotype in the group.

It is clear that the usual gene model, as given in the section for groups of size one, must be extended to describe not only *direct* effects due to genes specified by the genotype under consideration but also *associate* gene effects contributed by the other member of the group. Use of such an extended model, then, permits the accommodation of interactions between the two members of the group.

Although only alleles at one locus are considered, the genotypic value of ${}_{i_1j_1}d_{i_2j_2}$ is best described with a two-locus model as follows:

$$\begin{aligned} {}_{i_1j_1}d_{i_2j_2} = {}& {}_d a_{i_1} + {}_d a_{j_1} + {}_d \delta_{i_1j_1} + {}_a a_{i_2} + {}_a a_{j_2} + {}_a \delta_{i_2j_2} \\ & + {}_{da}(\alpha\alpha)_{i_1i_2} + {}_{da}(\alpha\alpha)_{i_1j_2} + {}_{da}(\alpha\alpha)_{j_1i_2} + {}_{da}(\alpha\alpha)_{j_1j_2} \\ & + {}_{da}(\alpha\delta)_{i_1i_2j_2} + {}_{da}(\alpha\delta)_{j_1i_2j_2} + {}_{da}(\delta\alpha)_{i_1j_1i_2} + {}_{da}(\delta\alpha)_{i_1j_1j_2} \\ & + {}_{da}(\delta\delta)_{i_1j_1i_2j_2}, \end{aligned}$$

where

$${}_d a_{i_1} = {}_{i\cdot}d_{\cdot\cdot} = \sum p_{j_1} p_{i_2} p_{j_2} ({}_{i_1j_1}d_{i_2j_2})$$
$$= \text{direct additive effect of allele } A_{i_1},$$

$${}_d \delta_{i_1j_1} = {}_{i_1j_1}d_{\cdot\cdot} - {}_d a_{i_1} - {}_d a_{j_1}$$
$$= \text{direct dominance effect of } A_{i_1}A_{j_1},$$

$${}_a a_{i_2} = {}_{\cdot\cdot}d_{i_2\cdot} = \sum p_{i_1} p_{j_1} p_{j_2} ({}_{i_1j_1}d_{i_2j_2})$$
$$= \text{associate additive effect of } A_{i_2} \text{ as measured on } A_{i_1}A_{j_1},$$

$${}_a \delta_{i_2j_2} = {}_{\cdot\cdot}d_{i_2j_2} - {}_a a_{i_2} - {}_a a_{j_2}$$
$$= \text{associate dominance effect of } A_{i_2}A_{j_2} \text{ as measured on } A_{i_1}A_{j_1},$$

$$_{da}(\alpha\alpha)_{i_1i_2} = {}_{i_1}.d_{i_2}. - {}_{d}\alpha_{i_1} - {}_{a}\alpha_{i_2}$$
$$= \text{additive} \times \text{additive interaction effect between direct allele } A_{i_1} \text{ and associate allele } A_{i_2},$$

$$_{da}(\alpha\delta)_{i_1i_2j_2} = {}_{i_1}.d_{i_2j_2} - {}_{d}\alpha_{i_1} - {}_{a}\alpha_{i_2} - {}_{a}\alpha_{j_2} - {}_{a}\delta_{i_2j_2} - {}_{da}(\alpha\alpha)_{i_1i_2} - {}_{da}(\alpha\alpha)_{i_1j_2}$$
$$= \text{additive} \times \text{dominance interaction effect between direct allele } A_{i_1} \text{ and associate genotype } A_{i_2}A_{j_2},$$

$$_{da}(\delta\alpha)_{i_1j_1i_2} = {}_{i_1j_1}d_{i_2}. - {}_{d}\alpha_{i_1} - {}_{d}\alpha_{j_1} - {}_{d}\delta_{i_1j_1} - {}_{a}\alpha_{i_2} - {}_{da}(\alpha\alpha)_{i_1i_2} - {}_{da}(\alpha\alpha)_{j_1i_2}$$
$$= \text{dominance} \times \text{additive interaction effect between direct genotype } A_{i_1}A_{j_1} \text{ and associate allele } A_{i_2}, \text{ and}$$

$$_{da}(\delta\delta)_{i_1j_1i_2j_2} = {}_{i_1j_1}d_{i_2j_2} - {}_{d}\alpha_{i_1} - {}_{d}\alpha_{j_1} - {}_{d}\delta_{i_1j_1} - {}_{a}\alpha_{i_2} - {}_{a}\alpha_{j_2} - {}_{a}\delta_{i_2j_2}$$
$$- {}_{da}(\alpha\alpha)_{i_1i_2} - {}_{da}(\alpha\alpha)_{i_1j_2} - {}_{da}(\alpha\alpha)_{j_1i_2} - {}_{da}(\alpha\alpha)_{j_1j_2} - {}_{da}(\alpha\delta)_{i_1i_2j_2}$$
$$- {}_{da}(\alpha\delta)_{j_1i_2j_2} - {}_{da}(\delta\alpha)_{i_1j_1i_2} - {}_{da}(\delta\alpha)_{i_1j_1j_2}$$
$$= \text{dominance} \times \text{dominance interaction effect between direct genotype } A_{i_1}A_{j_1} \text{ and associate genotype } A_{i_2}A_{j_2}.$$

The total genotypic variance for ${}_{i_1j_1}d_{i_2j_2}$ may be partitioned symbolically as

$$\sigma_G^2 = {}_{d}\sigma_A^2 + {}_{d}\sigma_D^2 + {}_{a}\sigma_A^2 + {}_{a}\sigma_D^2 + {}_{da}\sigma_{AA}^2 + {}_{da}\sigma_{AD}^2 + {}_{da}\sigma_{DA}^2 + {}_{da}\sigma_{DD}^2,$$

where
$$\sigma_G^2 = \sum p_{i_1}p_{j_1}p_{i_2}p_{j_2}({}_{i_1j_1}d_{i_2j_2})^2,$$
$${}_{d}\sigma_A^2 = 2\sum p_{i_1}({}_{d}\alpha_{i_1})^2,$$
$${}_{d}\sigma_D^2 = \sum p_{i_1}p_{j_1}({}_{d}\delta_{i_1j_1})^2,$$
$${}_{a}\sigma_A^2 = 2\sum p_{i_2}({}_{a}\alpha_{i_2})^2,$$
$${}_{a}\sigma_D^2 = \sum p_{i_2}p_{j_2}({}_{a}\delta_{i_2j_2})^2,$$
$${}_{da}\sigma_{AA}^2 = 4\sum p_{i_1}p_{i_2}[{}_{da}(\alpha\alpha)_{i_1i_2}]^2,$$
$${}_{da}\sigma_{AD}^2 = 2\sum p_{i_1}p_{i_2}p_{j_2}[{}_{da}(\alpha\delta)_{i_1i_2j_2}]^2,$$
$${}_{da}\sigma_{DA}^2 = 2\sum p_{i_1}p_{j_1}p_{i_2}[{}_{da}(\delta\alpha)_{i_1j_1i_2}]^2,$$
$${}_{da}\sigma_{DD}^2 = \sum p_{i_1}p_{j_1}p_{i_2}p_{j_2}[{}_{da}(\delta\delta)_{i_1j_1i_2j_2}]^2.$$

The gene model and genotypic variance components for ${}_{i_2j_2}d_{i_1j_1}$ are obtained from those given above for ${}_{i_1j_1}d_{i_2j_2}$ by the interchange of subscripts $i_1 \to i_2$ and $j_1 \to j_2$.

For prediction purposes, the following covariance between direct and associate effects must be defined:

$${}_{(da)}\sigma_A = 2\sum p_{i_1}({}_{d}\alpha_{i_1})({}_{a}\alpha_{i_1}).$$

(ii) *Consequences of Individual Selection*

The selection value for $A_{i_1}A_{j_1}$ when summed over all groups is

$$w_{i_1j_1} = 1 + (\bar{i}/\sigma)_{\text{ind.}}({}_{i_1j_1}d_{..}).$$

The expected gametic array produced by $A_{i_1}A_{j_1}$ is $\frac{1}{2}(A_{i_1}+A_{j_1})$. Therefore, the expected gametic array from all selected individuals is

$$\tfrac{1}{2}\sum p_{i_1}p_{j_1}\,w_{i_1j_1}(A_{i_1}+A_{j_1})$$

which equals

$$\sum p^1_{i_1}(A_{i_1}),$$

where

$$p^1_{i_1} = p_{i_1}[1+(i/\sigma)_{\text{ind.}}({}_d a_{i_1})].$$

Hence the new group population mean is

$$\mu_1 = \sum p^1_{i_1}p^1_{j_1}p^1_{i_2}p^1_{j_2}[\tfrac{1}{2}({}_{i_1j_1}d_{i_2j_2}+{}_{i_2j_2}d_{i_1j_1})],$$

or more simply

$$\begin{aligned}\mu_1 &= \sum p^1_{i_1}p^1_{j_1}p^1_{i_2}p^1_{j_2}({}_{i_1j_1}d_{i_2j_2})\\ &\simeq \sum p_{i_1}p_{j_1}p_{i_2}p_{j_2}\{1+(i/\sigma)_{\text{ind.}}[({}_d a_{i_1}+{}_d a_{j_1})+({}_d a_{i_2}+{}_d a_{j_2})]\}({}_{i_1j_1}d_{i_2j_2})\\ &= (i/\sigma)_{\text{ind.}}\{{}_d\sigma^2_A+{}_{(da)}\sigma_A\}.\end{aligned}$$

Since the original group population mean was coded to equal zero, the above value represents the increment change in mean due to one cycle of individual selection.

The important way in which this result differs from that obtained for groups of size one, is that when interactions among genotypes are taken into consideration the increment change is no longer a function of quantities that are necessarily non-negative. That is to say, if direct and associate effects for most genes are negatively related, the covariance ${}_{(da)}\sigma_A$ can be negative. Under the situation in which both members of the group are competing for the same environmental space, it is logical to assume that such a negative relationship often exists. Thus a gene which yields a positive direct advantage for the genotype containing it would tend to yield a negative associate stimulus to the competing member of the group.

If the covariance, ${}_{(da)}\sigma_A$, is negative and its absolute value is greater than that of ${}_d\sigma^2_A$, the very interesting situation occurs in which selection for individuals with greatest genotypic values results in a *decrease* of the progeny mean. Furthermore, since the increment change in group mean can be negative, continued positive selection can result in a final population mean which is different from its *maximum* potential value. These results are contrary to those obtained earlier for the situation in which associate effects are ignored (i.e. for groups of size one).

The suggestion may be made that selection for "general mixing ability" can solve the dilemma. In this case those genotypes which have highest yield when grown in association with all other genotypes have greatest selection values. However, again these selection values are a function of only direct gene effects and the results are similar to those of individual selection.

The question naturally arises, can a selection procedure be specified which ensures that positive selection will not lead to a deterioration of the population structure. More exactly, is it possible to develop selection procedures in which positive selection pressure invariably results in a non-negative increment change

in the mean, irrespective of the genetic relationship between direct and associate additive effects? The answer is unquestionably yes. In fact there are various selection procedures available, the most obvious of which is discussed in the next section. Other methods will be presented in subsequent papers of the series.

(iii) *Consequences of Group Selection*

This section considers consequences of selection based on group rather than individual performance. That is to say, the entire group is accepted or rejected on the basis of the average group mean. In this case, the selection value for $(A_{i_1}A_{j_1}, A_{i_2}A_{j_2})$ is

$$w_{i_1j_1,i_2j_2} = 1+(\bar{i}/\sigma)_{\text{gr.}}[\tfrac{1}{2}({}_{i_1j_1}d_{i_2j_2}+{}_{i_2j_2}d_{i_1j_1})],$$

where the subscript "gr." indicates that $\bar{i}$ and σ relate to group means.

Since the expected gametic array from such a group is $\frac{1}{4}(A_{i_1}+A_{j_1}+A_{i_2}+A_{j_2})$, the total expected gametic array following selection is

$$\tfrac{1}{4}\sum p_{i_1}p_{j_1}p_{i_2}p_{j_2}(w_{i_1j_1,i_2j_2})(A_{i_1}+A_{j_1}+A_{i_2}+A_{j_2})$$

which equals

$$\sum p^1_{i_1}(A_{i_1}),$$

where

$$p^1_{i_1} = p_{i_1}[1+(\bar{i}/\sigma)_{\text{gr.}}\tfrac{1}{2}({}_da_{i_1}+{}_aa_{i_1})].$$

Hence the increment change in group mean is

$$\begin{aligned}\Delta\mu &= \sum p^1_{i_1}p^1_{j_1}p^1_{i_2}p^1_{j_2}[\tfrac{1}{2}({}_{i_1j_1}d_{i_2j_2}+{}_{i_2j_2}d_{i_1j_1})]\\ &= \tfrac{1}{2}(\bar{i}/\sigma)_{\text{gr.}}[{}_d\sigma^2_A+2{}_{(da)}\sigma_A+{}_a\sigma^2_A].\end{aligned}$$

It is perhaps more obvious that this increment cannot be negative, irrespective of the relationships existing between direct and associate additive effects, when it is recast as the following sum of squares:

$$\Delta\mu = (\bar{i}/\sigma)_{\text{gr.}}[\sum p_{ik}({}_da_{ik}+{}_aa_{ik})^2].$$

This result demonstrates that transferring selection from an individual to a group basis automatically ensures that the population mean will not decrease due to positive selection.

(iv) *Numerical Example*

It is worth while to illustrate the points made concerning individual and group selection with a numerical example. For simplicity consider two equally frequent alleles which combine to give genotypes having the following values:

$$\begin{array}{lll} {}_{11}d_{11} = -2 & {}_{11}d_{12} = 0 & {}_{11}d_{22} = 6 \\ {}_{12}d_{11} = -2 & {}_{12}d_{12} = 0 & {}_{12}d_{22} = 6 \\ {}_{22}d_{11} = -6 & {}_{22}d_{12} = -4 & {}_{22}d_{22} = 2 \end{array}$$

In this case,

$$\begin{array}{lll} {}_d\sigma^2_A = 2 & {}_{(da)}\sigma_A = -4 & {}_a\sigma^2_A = 8 \\ {}_d\sigma^2_D = 1 & {}_{(da)}\sigma_D = -1 & {}_a\sigma^2_D = 1 \end{array}$$

The unselected population mean for groups of size two is zero. The increment change in the frequency of A_1 after one cycle of positive selection based on individual merit is

$$\Delta p_1 = \tfrac{1}{2}(\bar{i}/\sigma)_{\text{ind.}},$$

and the increment change in the population mean is *negative*, i.e.

$$\Delta\mu = -2(\bar{i}/\sigma)_{\text{ind.}}.$$

It can be shown that continued individual selection results in the fixation of the allele A_1 and hence the population mean steadily decreases from zero to the final value $\tfrac{1}{2}({}_{11}d_{11}+{}_{11}d_{11})$ which equals -2. This result is the worst of the two possible conditions of gene fixation.

On the other hand, group selection results in a negative increment change in the frequency of A_1, i.e.

$$\Delta p_1 = -\tfrac{1}{4}(\bar{i}/\sigma)_{\text{gr.}}.$$

The increment change in the population mean is *positive*, i.e.

$$\begin{aligned}\Delta\mu &= (\bar{i}/\sigma)_{\text{gr.}}\tfrac{1}{2}[2+2(-4)+8]\\ &= (\bar{i}/\sigma)_{\text{gr.}}\\ &> 0.\end{aligned}$$

Continued group selection results in fixation of A_2, and, hence, the population becomes homogeneous for groups of the constitution $(A_2A_2,\ A_2A_2)$, with mean $\tfrac{1}{2}({}_{22}d_{22}+{}_{22}d_{22})$, which equals 2. This represents the maximum potential for the random-mating population.

(c) *Groups of Size* n

(i) *Population Parameters*

The population of groups of size n is obtained from the n-way combinatorial product involving the base population genotypic array as follows:

$$\begin{aligned}&[\textstyle\sum p_ip_j(A_iA_j)]\times[\sum p_ip_j(A_iA_j)]\times\ \dots\ \times[\sum p_ip_j(A_iA_j)]\\ &\qquad = \textstyle\sum p_{i_1}p_{j_1}\dots p_{i_n}p_{j_n}(A_{i_1}A_{j_1},\ A_{i_2}A_{j_2},\ \dots,\ A_{i_n}A_{j_n}).\end{aligned}$$

The genotypic value of $A_{i_1}A_{j_1}$, as expressed in the n-tuple designated above, is denoted as ${}_{i_1j_1}d_{i_2j_2,\ \dots\ ,i_nj_n}$ and coded so that

$$\textstyle\sum p_{i_1}p_{j_1}\dots p_{i_n}p_{j_n}({}_{i_1j_1}d_{i_2j_2,\ \dots\ ,i_nj_n}) = 0.$$

This value may be represented as an extension of the model for groups of size two, i.e.

$$\begin{aligned}{}_{i_1j_1}d_{i_2j_2,\ \dots\ ,i_nj_n} = {}&{}_a\alpha_{i_1}+{}_a\alpha_{j_1}+{}_d\delta_{i_1j_1}+{}_a\alpha_{i_2}+{}_a\alpha_{j_2}+{}_a\delta_{i_2j_2}+\ \dots\\ &+{}_a\alpha_{i_n}+{}_a\alpha_{j_n}+{}_a\delta_{i_nj_n}+{}_{da}(\alpha\alpha)_{i_1j_2}+\ \dots\ .\end{aligned}$$

The entire model can be generated by expanding the product

$$(1+{}_d\alpha_{i_1}+{}_d\alpha_{j_1}+{}_d\delta_{i_1j_1})\Pi(1+{}_a\alpha_{i_k}+{}_a\alpha_{j_k}+{}_a\delta_{i_kj_k})$$

where in the expanded form,

(1) put the term "1" equal to zero,

(2) let $1\cdot({}_d\alpha_{i_1}) = {}_d\alpha_{i_1}$, $1\cdot({}_a\alpha_{i_k}) = {}_a\alpha_{i_k}$, and

(3) let $({}_d\alpha_{i_1})({}_a\alpha_{i_2}) = {}_{da}(\alpha\alpha)_{i_1i_2}$, $({}_a\alpha_{i_2})({}_a\delta_{i_3j_3}) = {}_{aa}(\alpha\delta)_{i_2i_3j_3}$, etc.

The total genotypic variance for ${}_{i_1j_1}d_{i_2j_2,\ldots,i_nj_n}$ may be partitioned in a manner which extends naturally from that in groups of size two, i.e.

$$\begin{aligned}\sigma_G^2 = {}&{}_d\sigma_A^2+{}_d\sigma_D^2+(n-1)({}_a\sigma_A^2+{}_{da}\sigma_{AA}^2+{}_{da}\sigma_{DA}^2)\\ &+(n-1)({}_a\sigma_D^2+{}_{da}\sigma_{AD}^2+{}_{da}\sigma_{DD}^2)+\cdots\\ &+\frac{(n-1)!}{(k_A)!(k_D)!(n-k_A-k_D-1)!}\Big[\underbrace{a\text{——}a}_{k_A+k_D}{}^{a}\sigma^2_{\underbrace{A\text{——}A}_{k_A}\,\underbrace{D\text{——}D}_{k_D}}+\underbrace{da\text{——}a}_{k_A+k_D}{}^{a}\sigma^2_{\underbrace{AA\text{——}A}_{k_A}\,\underbrace{D\text{——}D}_{k_D}}\\ &\qquad+\underbrace{da\text{——}a}_{k_A+k_D}{}^{a}\sigma^2_{\underbrace{DA\text{——}A}_{k_A}\,\underbrace{D\text{——}D}_{k_D}}\Big]+\cdots,\end{aligned}$$

where k_A = number of A subscripts in the variance component notation due to associate effects,

k_D = number of D subscripts in the variance component notation due to associate effects, and

n = number of individuals in the group.

Values of k_A and k_D can each range from zero to $n-1$, subject to the condition that their sum must not exceed $n-1$.

The entire model can be obtained by expanding the product

$$(1+{}_d\sigma_A^2+{}_d\sigma_D^2)(1+{}_a\sigma_A^2+{}_a\sigma_D^2)^{n-1},$$

where in the expansion

(1) put the term "1" equal to zero,

(2) let $({}_d\sigma_A^2)({}_a\sigma_A^2) = {}_{da}\sigma_{AA}^2$, and

(3) $\underbrace{({}_a\sigma_A^2)({}_a\sigma_D^2)\cdots({}_a\sigma_A^2)}_{t\text{ terms}} = {}^{a}\sigma^2_{\underbrace{a\text{——}a}_{t}\,\underbrace{AD\text{——}A}_{t}}$, etc.

(ii) *Consequences of Individual Selection*

The selection value for $A_{i_1}A_{j_1}$ when summed over all groups is

$$w_{i_1j_1} = 1+(i/\sigma)_{\text{ind.}}({}_{i_1j_1}d_{..,\text{—},..}).$$

The expected gametic array from selected individuals is $\sum p^1_{i_1}(A_{i_1})$, where

$$p^1_{i_1} = p_{i_1}[1+(i/\sigma)_{\text{ind.}}({}_d\alpha_{i_1})].$$

Hence the new group population mean is

$$\begin{aligned}\mu_1 &= \sum p^1_{i_1}p^1_{j_1}\cdots p^1_{i_n}p^1_{j_n}({}_{i_1j_1}d_{i_2j_2,\ldots,i_nj_n})\\ &= (i/\sigma)_{\text{ind.}}[{}_d\sigma^2_A+(n-1)_{(dd)}\sigma_A],\end{aligned}$$

which may be recast as

$$(i/\sigma)_{\text{ind.}}\{2\sum p_{i_1}({}_d\alpha_{i_1})[{}_d\alpha_{i_1}+(n-1)_a\alpha_{i_1}]\}.$$

The critical elements to be considered in this sum of cross-products are those represented by

$$[{}_d\alpha_{i_1}+(n-1)_a\alpha_{i_1}].$$

With groups of size one, the associate effect ${}_a\alpha_{i_1}$ disappears and the sum of cross-products becomes a sum of squares. This implies that the increment change in mean is non-negative and positive selection results in positive advance of the population mean.

It appears from the above formulation that as n increases associate effects may take on an increasingly dominant role in determining consequences of selection. For example, with groups of size two the increment change in population mean due to individual selection is

$$\Delta\mu = (i/\sigma)_{\text{ind.}}\{{}_d\sigma^2_A+{}_{(dd)}\sigma_A\},$$

which for two alleles at the locus, can be recast as

$$\Delta\mu = (i/\sigma)_{\text{ind.}}\{2(p_1/p_2)({}_d\alpha_1)({}_d\alpha_1+{}_a\alpha_1)\}.$$

Hence for $\Delta\mu$ to be negative, ${}_a\alpha_1$ must be of opposite sign to that of ${}_d\alpha_1$ and of greater absolute magnitude, i.e.

$$|{}_a\alpha_1| > |{}_d\alpha_1|.$$

With groups of size n, assuming two alleles at the locus, the increment change in population mean due to individual selection can be given in the same symbols as

$$\Delta\mu = (i/\sigma)_{\text{ind.}}\{2(p_1/p_2)({}_d\alpha_1)[{}_d\alpha_1+(n-1)_a\alpha_1]\},$$

where, for $\Delta\mu<0$, ${}_a\alpha_1$ must be of opposite sign to that of ${}_d\alpha_1$ but, in this instance, the magnitude of ${}_a\alpha_1$ need only be greater than $1/(n-1)$ times that of ${}_d\alpha_1$, i.e.

$$|{}_a\alpha_1| > \frac{1}{n-1}\left|{}_d\alpha_1\right|.$$

This implies that even for weakly competitive conditions a negative response to positive individual selection can occur.

In these arguments it must be understood that additive effects in groups of a given size need not be identical to those in groups of a different size. However, it is clear that as the number of effective members in the group increases, the range of possibilities increases in which a negative response to direct selection can occur.

(iii) *Consequences of Group Selection*

The selection value for the group $(A_{i_1}A_{j_1}, A_{i_2}A_{j_2}, \ldots, A_{i_n}A_{j_n})$ is

$$w_{i_1j_1,\ldots,i_nj_n} = 1+(\tilde{\imath}/\sigma)_{\mathrm{gr.}}[(1/n)({}_{i_1j_1}d_{i_2j_2,\ldots,i_nj_n}+\cdots+{}_{i_nj_n}d_{i_1j_1,\ldots,i_{n-1}j_{n-1}})].$$

The gametic array from selected individuals is $\sum p^1_{i_1}(A_{i_1})$, where

$$p^1_{i_1} = p_{i_1}\{1+(\tilde{\imath}/\sigma)_{\mathrm{gr.}}(1/n)[{}_d\alpha_{i_1}+(n-1)_a\alpha_{i_1}]\}.$$

The population mean is then

$$\mu_1 = (1/n)(\tilde{\imath}/\sigma)_{\mathrm{gr.}}\{{}_d\sigma^2_A+2(n-1)_{(da)}\sigma_A+(n-1)^2{}_a\sigma^2_A\}.$$

This expression can be recast as a sum of squares as follows:

$$\mu_1 = (1/n)(\tilde{\imath}/\sigma)_{\mathrm{gr.}}\{2\sum p_{i_1}[{}_d\alpha_{i_1}+(n-1)_a\alpha_{i_1}]^2\},$$

which demonstrates that it is a non-negative quantity. Thus, irrespective of the effective group size, non-negative genetic advance can be ensured by transferring the basis of selection from that of the individual to that of the group.

III. Discussion

In the most general terms, a study of the consequences of selection as it pertains to interacting genotypes requires a consideration of two populations of groups. That is to say, it is necessary to define not only the population of groups on which selection operates but also the population on which the effects of selection are to be measured. In the main part of this study these two populations are assumed to be the same, but clearly this need not be the case. The reason for the importance of distinguishing between the two populations is that the total effect of selection is measured as a sum of cross-products between certain effects of the gene model associated with the population in which selection takes place and the effects associated with the population used for evaluation. It has been shown that the most desirable selection procedures are those that convert this sum of cross-products, which can be negative, to a sum of squares, which, of course, cannot be negative.

In the present study it was demonstrated that among the classes of effects in the gene model (additive, dominance, and epistatic) only additive effects contribute to selection response. Hence the question as to whether or not a given selection procedure invariably produces desirable results depends on similarity of the *additive* component of the gene models for the populations on which selection operates and is evaluated. If selection utilizes and is evaluated with exactly the same additive model, the sum of cross-products becomes a sum of squares and the change in mean cannot be negative. Continuation of this balanced selection procedure results in the most desirable final population.

It is clear that if group selection operates on and is evaluated using the same population, it satisfies the necessary requirements for balance. This is true because group selection operates on both direct and associate components and hence on the same total model as that with which it is evaluated. On the other hand, even if individual selection operates on and is evaluated using the same population, it *cannot* satisfy these requirements. This is true because it operates on only the direct additive effects and is evaluated with direct and associate effects; hence it is an unbalanced selection procedure. Likewise, a selection method which utilizes only the associate additive component cannot be a balanced precedure. More importantly, the same argument holds for situations in which selection of any sort operates on groups of a given size and is evaluated for groups of a different size. For example, in pasture breeding there is no guarantee that positive selection among spaced plants will not elicit a negative response when evaluated under sward conditions.

Although it is not entirely appropriate to apply the unordered group model to the barley studies of Wiebe, Petr, and Stevens (1963), it is sufficiently accurate to describe the basic problem in which a given genotype may express different values when placed in groups of different genotypes. The numerical example given earlier was constructed to illustrate this phenomenon. Hence, as with the barley studies, positive selection in the numerical example results in a negative response, and conversely, selection of the poorest genotypes results in an increase of the population mean. This pattern of response is basically due to the fact that a negative relationship between direct and associate effects was built into the example. However, in the real world such negative relationships need not always exist and knowledge as to whether or not they do exist is seldom available. Hence, even in strongly competitive situations, negative individual selection cannot be seriously recommended as a suitable procedure.

A solution to the dilemma is given in this study. It consists of developing a theory which accommodates the phenomenon in question and then with the help of this theory identifying those selection procedures which ensure non-negative responses in the population mean. It is shown in the present study that group selection satisfies the necessary requirements. In subsequent papers, other more specialized selection methods will be discussed which also invariably satisfy the requirements and which may be more efficient than group selection.

IV. References

Allard, R. W. (1961).—Relationship between genetic diversity and consistency of performance in different environments. *Crop Sci.* **1**, 127–33.

Hamilton, W. D. (1964).—The genetical evolution of social behaviour. I. *J. Theoret. Biol.* **7**, 1–16.

Lewontin, R. C. (1965).—Selection in and of populations. In "Ideas in Modern Biology". (Ed. J. A. Moore.) (Proc. 16th Int. Congr. Zool.) Vol. 6. pp. 299–311. (Natural History Press: Garden City, New York.)

Simmonds, N. W. (1962).—Variability in crop plants, its use and conservation. *Biol. Rev.* **37**, 422–65.

Suneson, C. A. (1956).—An evolutionary plant breeding method. *Agron. J.* **48**, 188–91.

Wiebe, G. A., Petr, F. C., and Stevens, H. (1963).—Interplant competition between barley genotypes. In "Statistical Genetics and Plant Breeding". (Eds. W. D. Hanson and H. F. Robinson.) pp. 546–55. (National Academy of Science: National Research Council Publ. No. 982.)