

18

Associative Effects: Competition, Social Interactions, Group and Kin Selection

These findings ... support the writer's view that competitive ability should be accepted as it stands as a genetic character, simple or aggregate, a view of great importance in the discussion to follow. — Sakia (1955)

Draft Version 13 July 2010.

This chapter weaves together several seemingly unrelated, but nevertheless important, topics: competition; altruism and other social behaviors; traits defined by group, rather than individual, behavior; maternal effects; group and kin selection. The connection between all of these topics is the notion that the genotype (and hence phenotype) of one individual may influence the trait value of another. In this sense, the “environmental” component of the phenotype of a focal individual may itself have some heritable component (based on the contribution from some other individual), allowing some part of the environmental component to evolve along with the focal trait. In such settings, the phenotype of a focal individual consists of two components: **direct effects** from the focal individual and **associative** (or **as-sociate**) **effects** contributed from other individuals within the group. A critical implication of this distinction is that the breeding value of an individual contains a component for direct effects that appear in its phenotype (and hence can be influenced by individual selection) and a component for associative effects that only appear in the phenotypes of *other* group members. For selection to effectively access the associative contribution generally requires either that (i) interactions occur among kin (**kin selection**), or (ii) selection is based on some combination of both individual and group values (**multilevel selection**). In the extreme, **group selection** occurs when all of the weight is placed on between-group differences. Note that multilevel selection is a generalization of the family selection index (Chapter 17) to more general groups.

The framework for dealing with these issues has been laid out in a series of classic, but largely ignored, papers by Griffing (1967, 1968a, 1968b, 1969, 1976a, 1976b, 1977), who introduced associative effects. There are also roots extending to classic work on maternal effects based on trait phenotype (Falconer 1965) or on an unmeasured material value (Willham 1963), as well as to the foundations of the study of social evolution (Hamilton 1963, 1964a,b). There are two modeling approaches for dealing with associative effects: **trait-based** and **variance-component based**. Trait-based approaches (Moore et al. 1997) have their roots in univariate (Falconer 1965) and multivariate (Kirkpatrick and Lande 1989, Lande and Kirkpatrick 1990) models of response under maternal effects. As their name implies, trait-based approaches assume that we know the particular traits in group members that influence the phenotype of the focal individual. This approach is best handled in a multivariate framework, so we delay discussion until Chapter 37. The variance-component approach also has roots in maternal-effects models (Willham 1963) wherein a general (but unmeasured) maternal performance value influences the phenotype of the focal individual. Using BLUP, we can estimate the genetic variance of the associative effects (as well as its covariance with the

direct effects). Somewhat counter-intuitively, variance-component based methods (where the actual traits that generate the associative effects are unspecified) is empirically more powerful than trait-based methods. The reason is that we can estimate this unspecified total contribution directly, while if traits that influence associative effects are left out of the model, this can introduce errors. McGlothlin and Brodie (2009) show the congruence between these methods, which is also examined in detail in Chapter 37.

Traits whose phenotypes are determined, in part, by interactions with other individuals have important roles in breeding and evolution. In breeding, we are often more interested in the performance of a group rather than an individual. For example, standard poultry husbandry is to keep several females together in a cage, with total egg production per cage being the key quantity of interest. In the extreme, an aggressive female may kill all her cage-mates, and in less extreme cases may largely dominate feeding, resulting in an individual benefit at the expense of the group. Hence, individual selection may result in a *decrease* in group performance, in which case eggs per cage would decline. The issue here is that individual selection cannot often effectively utilize any genetic variation in associative effects to move the mean associative value in the direction favored by the breeder. The same concerns have long been raised in evolutionary biology, in particular to account for the evolution of altruistic traits (such as alarm calls in birds) that are expected to decrease individual fitness, yet still have evolved. There is a very rich, and stormy, evolutionary literature on the importance (or lack thereof) of selection based on group attributes. The general view in evolutionary biology has often been to invoke group selection arguments only as a last, desperate resort when all individual selection arguments fail (e.g., Williams 1966). As we will see, much of the debate regarding group- vs. kin-selection is misplaced, as they are essentially manifestations of the same general process.

Our treatment starts with a formal definition of direct and associative effects, including the powerful concept of the total breeding value of a trait (which requires measurements of group members). Next, we show how the presence of associative effects influences selection. One key result is that when the breeding values for direct and associative effects are negatively correlated, individual selection can give a reversed response. Conversely, group selection (even when group members are unrelated) always results in an expected positive response, but can be very ineffective when associative effects are small. We then examine selection based on an index of both individual and group information, including the optimal weighting for maximal response. A key innovation is the use of BLUP/REML methodology to estimate the direct and associative effects of individuals along with their variance components, and we examine this in detail. Finally, we conclude by (briefly) applying these results to some of the debates on group and kin selection in evolutionary biology. Our goal in this last section is not to extensively review this literature, which is often contradictory and at times was driven more by verbal models than detailed analysis. Rather, it is to show how the problem of selection based on group attributes can be easily placed in a quantitative-genetics framework.

DIRECT VERSUS ASSOCIATIVE EFFECTS

All organisms interact with their environment, and a very significant fraction of that environment is biological. In particular, interactions with conspecifics through competition, cooperation, parental care, or other social interactions can constitute an important part of the environment that an individual experiences and, in turn, this can influence trait values. Further, this “environment” may contain heritable components and coevolve with the trait of interest. The classic example of this is competition and we briefly consider this first.

Early Models of Competition

It has long been appreciated by breeders that competition among plants within a plot has a significant impact on important agricultural traits such as yield. While a particular genotype when grown in isolation may have high yield, when grown in a group, its competitive effects on other members within its group could actually result in lower plot yield. Yield (and other traits) of a particular plant in a plot is therefore a function of two components. First, the individuals' genotype has a direct influence on its ability to garner resources such as light, water, and nutrients. Second, that genotype influences others around it, by competing for limiting resources. Other plants in the plot also compete, and these in turn influence the yield of the focal individual. One might expect that plants that are very successful at garnering resources have positive direct effects but negative associative effects on nearby individuals. Thus, a plot of high-competing genotypes can have a low yield, as the positive direct effects for any particular plant are more than countered by negative associative effects from being surrounded by superior competitors.

Example 18.1. This point was made in a classic paper by Weibe et al (1976), who examined yield in mixed- versus single-genotype plots of barley. They observed that genotypes which yielded well in mixed stands had poorer yield in pure stands, while those genotypes that did poorly in mixed stands had the highest yield in pure stands. In our framework, we could imagine that lines which do well in mixed stands have both high direct effects and high negative associative effects, suppressing the phenotypes of their neighbors. When grown in a pure stand, the high negative associative effects suppress plot yield. Conversely, lines that perform poorly in mixed stands might have low direct effects but high positive associative effects, so that the phenotypes of their neighbors are enhanced (or at least not hindered). When grown as a pure stand, these high positive associative effects more than compensate for the low direct effects, increasing yield.

A historically important paper on the evolution of competition was Sakia (1955), who noted that competition, like yield or height, is a genetic trait and hence can potentially evolve. Following Sakia, a number of workers developed single-locus population genetic models to examine the evolution of competition (Schutz et al. 1968, Schutz and Usanis 1969, Allard and Adams 1969, Cockerham and Burrows 1971, Cockerham et al. 1972). These models all used simple ecological models of competition among a series of fixed types (here, all possible single-locus genotypes). While interesting, this class of models does not easily generalize beyond one locus. Griffing (1967) made the important extension of Sakia's idea by replacing a single-locus genotype with direct and associate values that are quantitative traits, consisting of breeding and residual values. Placed in this framework, such traits can potentially evolve and can also have their variance components estimated, allowing associative effects to be exploited by using appropriate selection designs.

Direct and Associative Effects

A simple example will introduce Griffing's idea. As shown in Figure 18.1A, consider a group of four individuals. Our focal individual is 1, and its phenotype (for our trait of interest) is determined by its own intrinsic value $P_{d,1}$ (with the subscript d indicating the direct effect) plus associative effects $P_{s,2}$, $P_{s,3}$, and $P_{s,4}$, contributed by other group members.

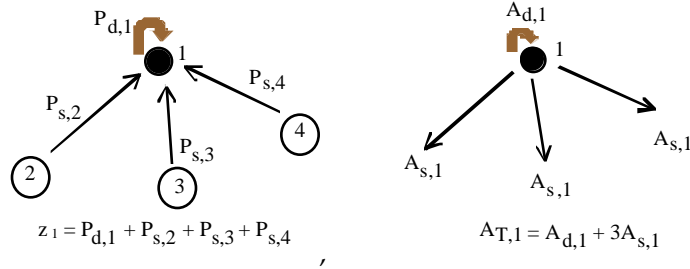


Figure 18.1. **A: Left:** The phenotypic value z_1 of the focal individual is the sum of its direct phenotypic effect ($P_{d,1}$) plus the associative effects $P_{s,2}$, $P_{s,3}$, $P_{s,4}$ of (in this case) the three other members in its group. **B: Right:** The total breeding value ($A_{T,1}$) of individual one is its direct breeding value $A_{d,1}$ plus the total contribution of the associative effect breeding value $A_{s,1}$ to the three members of its group. Only part of A_T (namely A_d) is embedded within its own phenotypic value. The remaining part of A_t , namely its associative component, is only expressed in the phenotypes of *other* group members.

Associative/associate effects are also referred to in the literature as **indirect genetic effects** (IGEs) (Wolf et al. 1998), or **social effects** (Bijma et al. 2007a,b). We use the subscript s (indicating social effects) to denote these, as the other local choice a might create some confusion given we use A for breeding values. In our discussion, we will use associate, associative, and social effects interchangeably. Note that the $P_{s,i}$ do not necessarily correspond to the phenotypes for the trait on interest in the other group members, but rather represent the contribution from these members to the phenotype of the focal individual — this contribution from fellow group members is part of the environment experienced by the focal individual. More generally, for a group of size n of equally-interacting individuals, the resulting phenotype z_i for individual i becomes

$$z_i = P_{d,i} + \sum_{j \neq i} P_{s,j} \quad (18.1a)$$

where the sum has $n - 1$ terms. Each of these components can be decomposed into a breeding value A plus a residual component E (containing environmental effects plus any nonadditive genetic variance), giving

$$z_i = \mu + (A_{d,i} + E_{d,i}) + \sum_{j \neq i} (A_{s,j} + E_{s,j}) \quad (18.1b)$$

We can write this compactly as

$$z_i = \mu + A_{d,i} + \sum_{j \neq i} A_{s,j} + e_i, \quad \text{where} \quad e_i = E_{d,i} + \sum_{j \neq i} E_{s,j} \quad (18.1c)$$

Since the environmental values have expected value zero, the mean phenotypic value in the group is just

$$\mu_z = \mu_{A_d} + (n - 1)\mu_{A_s} \quad (18.1d)$$

Further, the change in the mean trait value within a group following selection is

$$\Delta\mu_z = \Delta\mu_{A_d} + (n - 1)\Delta\mu_{A_s} \quad (18.1e)$$

which decomposes the changes in trait value into contributions from changes (responses) in direct and social values. This equation foreshadows individual versus group selection.

Individual selection targets the direct effect and results in a favorable change in μ_{A_d} . If the direct and social breeding values are correlated within an individual ($\sigma(A_d, A_s) \neq 0$), then individual selection can also change μ_{A_s} , but not necessarily in a favorable direction. Indeed, as Example 18.4 shows, an increase in μ_{A_d} under individual selection can be more than countered by an unfavorable change in μ_{A_s} , resulting in the mean phenotype changing in an unfavorable direction. Direct selection on μ_{A_s} requires either group selection and/or having relatives within the group. All of these points will be expanded upon below. Finally, our focus is entirely on additive genetic effects, as most of the theory has been developed under this assumption. Attempts to include non-additive variance are given by Gallais (1976) and Wright (1986).

Animal Well-being and the Improvement of the Heritable Social Environment

In high-intensity agricultural systems, competition has a strong effect on yield and other traits. Production animals in such environments face significant stress that not only impacts system production but also their own well-being. As reviewed by Muir (1998), animal well being is becoming an increasingly important aspect of animal production. Muir suggests that social aspects such as aggression, fighting, and sharing of common resources are all potential targets of selection, and responses in these traits (for less aggression and more sharing) improves both welfare and production. Specifically, Muir suggested improving welfare by selecting for improved mean social environment by choosing individual with favorable A_s values for the trait(s) of interest. Further, for a number of species (such as certain fishes), domestication has proved somewhat problematic due to the tendency for cannibalism and lesser forms of aggression when grown under production conditions. Again, these are aspects of the group environment and can respond favorably to an appropriate selection design, provided there is a heritable component of P_s (i.e., that $\sigma_{A_s}^2 > 0$).

What Do We Mean by Group?

Given that we use the term “group” extensively in this chapter, a more formal definition is required. Our focus here is on traits whose values are influenced by interactions with others. The set of individuals that interacts with the focal individual constitutes the unit we will call a group. In some breeding setting, this may be straightforward, such as the number of animals in a pen or cage. However, in other breeding setting, such as cattle in a very large feedlot, only some subset of the individuals within the feedlot likely interact with the focal individual. Hence, group size may be much smaller than the number of individuals physically confined to some space. Likewise, individuals may be part of different groups for different traits, especially if those traits are expressed at different times during development. The same is true on a grander scale in natural populations. The key issue with traits influenced by interactions is that phenotypes of the group members provide some information on part of the breeding value of the focal individual, that dealing with associative effects, that is simply not provided from the phenotype of the focal individual. To exploit this additional heritable variation (when it exists), interactions with relatives and/or selection that puts at least some weight on group value is essential.

The second feature about groups is their formation and reproduction, an issue especially important under differential propagation of groups (i.e., group selection). Here we are assuming a situation akin to our analysis of family selection (Chapter 17) in that while group information may be used to select individuals to form the next generation, these individuals are then mated at random. In the group selection literature, this is referred to as a **migrant pool model** (Levins 1970, Wade 1978). Such a structure only allows changes in breeding values to propagate to the next generation. In settings where entire groups are propagate as a unit (the **propagule pool model**, Wade 1978) the potential for non-additive variance to

contributed to the between group variance exists.

Trait- Versus Variance-Component Based Models

A brief comment is on order expanding upon our earlier remarks on trait vs. variance component based modeling. The original trait-based model of associative effects was Falconer's (1965) model for litter size in mice (Equation 11.20),

$$z_i = G_i + e_i + m \cdot z_{mo,i}$$

where G_i is the direct breeding value for litter size, while the associative effect is a function of the litter size of its mother ($z_{mo,i}$). Building on this idea, Moore et al. (1997) and Wolf et al. (1998) suggest the model that the value for trait i also depends upon the value z'_j of trait j (which may, or may not, equal i) in an interacting individual is given by

$$z_i = A_i + e_i + \psi_{ij} z'_j = A_i + e_i + \psi_{ij} A'_j + \psi_{ij} E'_j \quad (18.2)$$

where ψ_{ij} (following Kirkpatrick and Lande 1989) is the multivariate extension of Falconer's m . Under this model, breeding value for the associative effect is $\psi_{ij} A'_{s,j}$. This class of models leads to very interesting behavior, such as feedback loops that significantly modify Equation 18.2. Figure 18.2 illustrates this difference in modeling, and Chapter 37 explores trait-based models in some detail.

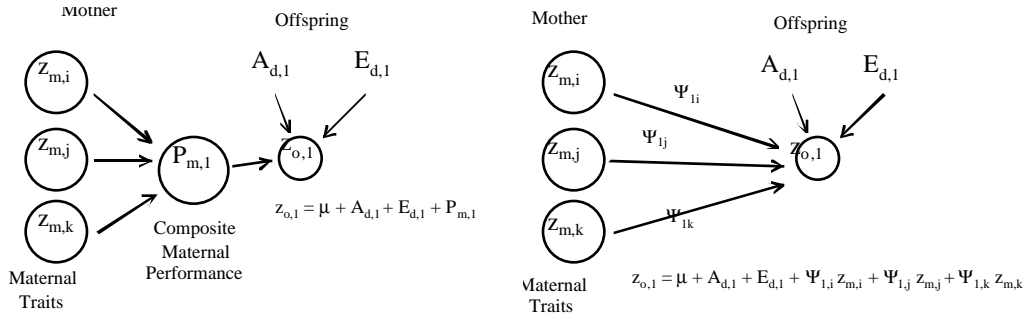


Figure 18.2. The difference between trait and variance-component based models. Here, the phenotypic value $z_{o,1}$ of a trait (which we label one) in an offspring is a function of maternal phenotype. We suppose that there are three maternal traits (i, j, k) whose phenotypes influence the offspring value. **A: Left:** Under a variance-component based approach, we ignore all the maternal trait values and simply estimate a single maternal performance value $P_{m,1}$ that directly influences the trait value in the offspring, so that the model becomes $z_{o,1} = \mu + A_{d,1} + E_{d,1} + P_{m,1}$, where $A_{d,1}$ is the trait breeding value in the offspring, $E_{d,1}$ its environmental value, and $P_{m,1} = A_{s,1} + E_{s,1}$ can be decomposed into its social breeding value on trait one plus a residual. **B: Right:** Under a trait-based models, if we know all of the maternal traits whose phenotypes influence trait one in the offspring, then we directly incorporate these, along with their regression coefficients $\Psi_{1,i}$ showing how these maternal phenotypes translate into offspring trait value. Here, $z_{o,1} = \mu + A_{d,1} + E_{d,1} + \Psi_{1,i} z_{m,i} + \Psi_{1,j} z_{m,j} + \Psi_{1,k} z_{m,k}$. The advantage of trait-based models is that they are required to following the joint evolution of traits 1 and i, j, k . Their drawback is that we have to know all of the important maternal traits. Conversely, under a variance-component method, all of the maternal phenotypes are conveniently collapsed into a single value, whose breeding value we can estimate from an appropriate design (detailed below).

The Total Breeding Value (TBV) and T^2

Given that an individual contains breeding values for both direct and social effects, what is their breeding value contribution to the next generation? We can directly see this from

Equation 18.1d, as the contribution to the population mean from individual 1 from its direct breeding value A_{d_1} plus its contribution to the $(n - 1)$ other individuals in its group through its associative effects breeding value A_{s_1} (Figure 18.1b). Based on this observation, Bijma et al. (2007a) define the **total breeding value** (or **TBV**), A_T , of a trait from an individual measured in a group of size n as the sum of its direct effect plus the total associative effects over all group members, or

$$A_{T_i} = A_{d_i} + (n - 1)A_{s_i} \quad (18.3)$$

Moore et al. (1997) introduced a similar measure for trait-based models. Noting that the mean of the population is simply the mean breeding value allows Equation 18.3 to recover Equation 18.1d. The critical observation is that when associative effects are present, the total breeding value of an individual contains components that are *not* expressed in its own phenotype, but rather only in the phenotypes of other individuals with which it interacts.

Example 18.2. Consider a trait in a group of four (unrelated) individuals, where (for illustrative purposes) we assume no environmental values so that $P_d = A_d$ and $P_s = A_s$. The population mean is 20, and the four group members have the following breeding values for direct, associative, and total effects:

Individual	A_d	A_s	A_T	$\sum_{j \neq i} A_{s_j}$	z
1	9	-4	-3	4	33
2	5	-1	2	1	26
3	-6	2	0	-2	12
4	-8	3	1	-3	9

Since $n = 4$, $A_T = A_d + 3A_s$. The sum $\sum_{j \neq i} A_{s_j}$ represents the contribution of the associative effects of the other three individuals to i 's value. For example, for individual 1, the contributions from individuals 2 through 4 is $-1 + 2 + 3 = 4$. From Equation 18.1c, the phenotypic value we would observe is

$$z_i = 20 + A_{d_i} + \sum_{j \neq i} A_{s_j}$$

Individual one has the largest direct effect (9) and the largest observed trait value (33). This individual also has the most unfavorable associative value (-4), and the smallest total breeding value (-3). Conversely, it has the largest contribution (4) to its trait value from the associative effects of the other group members. Its high trait value is due to this combination of a high direct effect and a high contribution from the associative effects of the other group members. Its unfavorable associative effects do not appear in its own phenotype, but rather are expressed in the trait values of the other group members. As a result, its own phenotypic value is a poor predictor of A_T .

If the next generation is formed by crossing the two individuals (1 and 2) with the largest trait values, the expected offspring mean is $20 + (-3+2)/2 = 19.5$, the mean plus their average total breeding values. Although the two largest individuals were chosen, the population mean *decreases*. Conversely, crossing the two smallest individuals gives an expected offspring mean of $20 + (0+1)/2 = 20.5$, increasing the mean. While the two smallest individuals have the smallest direct effects, they also have the most favorable associative effects, and hence give a more favorable response. The greatest expected response occurs by crossing the two individuals (2 and 4) with the largest total breeding values, for an expected mean of $20 + (2+1)/2 = 21.5$.

The covariance between an individual's phenotype and total breeding value is

$$\begin{aligned}\sigma(z_i, A_{T_i}) &= \sigma\left(\mu + A_{d_i} + \sum_{j \neq i} A_{s_j} + e_i, A_{d_i} + (n-1)A_{s_i}\right) \\ &= \sigma\left(A_{d_i}, A_{d_i} + (n-1)A_{s_i}\right) + \sum_{j \neq i} \sigma\left(A_{s_j}, A_{d_i} + (n-1)A_{s_i}\right)\end{aligned}\quad (18.4a)$$

For now, we assume unrelated group members, in which case the covariances in the summation are all zero, giving

$$\sigma(z, A_T) = \sigma^2(A_d) + (n-1)\sigma(A_d, A_s) \quad (18.4b)$$

If the direct and associative effects are uncorrelated, this reduces to our standard result that the covariance between an individual's phenotype and breeding value is just the additive genetic variance (in this case, of direct effects). By contrast, the variance of the total breeding value becomes

$$\begin{aligned}\sigma^2(A_T) &= \sigma^2[A_d + (n-1)A_s] \\ &= \sigma^2(A_d) + 2(n-1)\sigma(A_d, A_s) + (n-1)^2\sigma^2(A_s)\end{aligned}\quad (18.4c)$$

$$= \sigma(z, A_T) + (n-1)[2\sigma(A_d, A_s) + (n-1)\sigma^2(A_s)] \quad (18.4d)$$

Equation 18.4d shows that the covariance between total breeding value and phenotype is different from the variance in total breeding value. This reflects the fact that the associative effects of an individual do not influence its own phenotype. Note from Equation 18.4c that the contribution to the total additive variance from heritable social effects, $2(n-1)\sigma(A_d, A_s) + (n-1)^2\sigma^2(A_s)$, can be considerable. Further, note that the variance in total breeding value is a function of group size n , but is independent of the relationships among group members.

Now consider the phenotypic variance,

$$\sigma_z^2 = \sigma^2\left(P_{d_i} + \sum_{j \neq i} P_{s_j}\right). \quad (18.5a)$$

Assuming (for now) that the group members are unrelated, so that $\sigma(P_{d_i}, P_{s_j}) = 0$. For a group of size n Equation 18.5a reduces to

$$\sigma_z^2 = \sigma^2(P_d) + (n-1)\sigma^2(P_s) \quad (18.5b)$$

$$= \sigma^2(A_d) + (n-1)\sigma^2(A_s) + \sigma^2(E_d) + (n-1)\sigma^2(E_s) \quad (18.5c)$$

$$= \sigma^2(A_d) + (n-1)\sigma^2(A_s) + \sigma^2(e) \quad (18.5d)$$

where e is given by Equation 18.1c. With the phenotypic variance in hand, we can define the heritability of the direct and associative effects as

$$h_d^2 = \frac{\sigma^2(A_d)}{\sigma_z^2}, \quad \text{and} \quad h_s^2 = \frac{\sigma^2(A_s)}{\sigma_z^2} \quad (18.6a)$$

The careful reader will note that there is a different, but perhaps more natural, definition of these two heritabilities. Equation 18.6a standardizes the genetic variances with respect to the total trait variance, but one could also standardize them with respect to the variance of direct and associative effects, e.g.,

$$h_d'^2 = \frac{\sigma^2(A_d)}{\sigma^2(P_d)}, \quad \text{and} \quad h_s'^2 = \frac{\sigma^2(A_s)}{\sigma^2(P_s)} \quad (18.6b)$$

We use a prime to distinguish these from the heritabilities scaled to total trait variance (σ_z^2 vs. $\sigma_{P_x}^2$). While heritabilities scaled by σ_z^2 (i.e., Equation 18.6a) are the most widespread in the literature, there are some advantages to scaling then by $\sigma_{P_x}^2$. On this scale, the heritabilities measure the fraction of additive genetic variation in the actual effect (direct or associate) itself, rather than in the trait value. Further, $h_x'^2$ is independent of the group size (provided that A_s does not change with group size), as $\sigma_{P_x}^2$ is independent of n , while σ_z^2 is a function on n .

In keeping with Equation 18.6a, we can similarly define the “heritability” of the total breeding value as

$$T^2 = \frac{\sigma^2(A_T)}{\sigma_z^2} \quad (18.7)$$

as suggested by Bijma et al. (2007a). The reason we have used T^2 rather than h_T^2 is that, unlike heritabilities, T^2 can exceed one. This can happen because $\sigma^2(A_T)$ contains additional terms not found in σ_z^2 , as the associative effects of an individual influence others in the group, rather than the individual in which they reside. To see this, assume that the environmental effects are all zero, so that we can focus on differences in the genetic variance components. From Equations 18.4c and 18.5c,

$$\begin{aligned} \sigma^2(A_T) - \sigma_z^2 &= 2(n-1)\sigma(A_d, A_s) + (n-1)^2\sigma^2(A_s) - (n-1)\sigma^2(A_s) \\ &= (n-1) [2\sigma(A_d, A_s) + (n-2)\sigma^2(A_s)] \end{aligned} \quad (18.8)$$

If this difference exceeds the contribution σ_e^2 from environmental/residual effects, $T^2 > 1$.

Example 18.3. Consider a trait in a group of 10 unrelated individuals, with $\sigma_{P_d}^2 = 10$, $\sigma_{P_s}^2 = 1$, and both direct and associative effects have modest heritabilities measured on the scale of the effect themselves ($h_d'^2 = 0.4$, $h_s'^2 = 0.3$). To simplify matters, assume $\sigma(A_d, A_s) = 0$. Applying Equation 18.5b, the resulting phenotypic variance is

$$\sigma_z^2 = \sigma_{P_d}^2 + 9 \cdot \sigma_{P_s}^2 = 10 + 9 \cdot 1 = 19$$

From Equation 18.4c, the variance in total breeding value becomes

$$\sigma_{A_T}^2 = \sigma_{A_d}^2 + 9^2 \cdot \sigma_{A_s}^2 = h_d'^2 \sigma_{P_d}^2 + 9^2 \cdot h_s'^2 \sigma_{P_s}^2 = 4 + 81 \cdot 0.3 = 28.3,$$

giving $T^2 = 28.3/19 = 1.57$.

A real world example of large potential differences in $h_d'^2$ versus T^2 is survival days in chickens (Bijma et al. 2007b). Ignoring associative effects gives a heritability $h_d'^2$ of 6.7%, while when using a mixed model that incorporates associative effects (detailed later in the chapter), the estimate of T^2 was 20%, a threefold increase. Hence, under the conditions in the study, roughly two-thirds of the heritable variation in the trait arises from interactions between individuals and is thus hidden from standard analyses which ignore these. As discussed below, this component is only fully accessible under individual selection if the group includes relatives.

Notice that $\sigma(A_d, A_s)$ and $\sigma^2(A_s)$ are scaled by $(n-1)$ and $(n-1)^2$, respectively, in $\sigma^2(A_T)$. Hence, with even modest group sizes, small values of these two components can still have a very significant impact. Some of the early papers reporting estimates of these

two quantities ignored this scaling, and hence tended to down-play the importance of social interactions (Chen et al. 2006, Van Vleck et al. 2007).

A_s as a Function of Group Size

As the careful reader will have noted, the direct effect A_d is independent of group size, while the social effect A_s potentially changes over group size. Suppose a genotype has breeding value for social effects of 10 when measured in groups of size four. Does this change with group size and, if so, how? This is an empirical issue, and one can frame it in a $G \times E$ setting – the different environments here are group size, and if A_s shows $G \times E$, it changes over group size. Two simple scenarios bracket the possible changes. First, suppose that an individual eats 500 grams of food daily. In a group with fixed food size, the associative effects of this individual are to remove 500 grams from the total food supply each day. Hence, in a group of size n , $P_{s_i} = -500/(n-1)$, while its *total* associative effect is the sum over all group members, $(n-1)P_{s_i} = -500$. Here the total associative effect remains unchanged, while the individual social effect on any group member shows a **dilution** with increasing group size. Alternatively, consider a large tree whose associative effects results from its shading of any individuals under its canopy. In such a case, the associative effect shows no dilution with group size. Bijma et al. (2007a) note that alarm calls are also expected to show no dilution with group size.

More generally, we have been assuming that all group members experience the same social effect from a conspecific (i.e., all get $P_{s,i}$ from individual i). However, one can imagine settings where $P_{s,i}$ is some base effect, but its effect on specific individuals depends on their distance (e.g., Muir 2005, Cappa and Cantet 2008) or the total time that they interact (Cantet and Cappa 2008).

Clearly, a developing research area is the further refining of social effects and how they change over group size. Some initial insight is provided by Hadfield and Wilson (2007) and Bijma (2010b). Hadfield and Wilson assumed a simple regression model,

$$P_{s_i,n} = P_{sb_i} + \frac{1}{n-1} P_{sr_i} \quad (18.9)$$

with the value for social effect in a group of size n a function of two components: a base (or intercept) value P_{sb_i} plus a rate of change P_{sr_i} . Note that the resulting total sum over the $n-1$ group members becomes $(n-1)P_{sb_i} + P_{sr_i}$, so that P_{sr} is the constant contribution, while that from P_{sb} scales with group size.

Bijma (2010b) suggested a related model,

$$P_{s_i,n} = \frac{1}{(n-1)^d} P_{s_i,2} \quad (18.10a)$$

which expresses all group social values as a function of the value for a group of size two plus a **dilution fraction** d (assumed to be the same over all genotypes). Bijma's model is a bit more tractable, while the Hadfield-Wilson model is more general. When $d = 1$ and $P_{sp_i} = 0$, the two models are equivalent. Under the Bijma model, the total phenotypic variance becomes

$$\sigma_{z,n}^2 = \sigma^2(P_d) + (n-1)^{1-2d} \sigma^2(P_{s,2}) \quad (18.10b)$$

Phenotypic variance increases with n for $d < 1/2$, remains constant for $d = 1/2$ and decreases with n for $d > 1/2$. Assuming that breeding values are diluted in the same fashion as phenotypic effects,

$$A_{s,n} = \frac{A_{s,2}}{(n-1)^d} \quad \text{and} \quad \sigma^2(A_{s,n}) = \frac{\sigma^2(A_{s,2})}{(n-1)^{2d}} \quad (18.10c)$$

Substituting into Equation 18.4c gives the total additive variance for a group of size n as

$$\sigma^2(A_{T,n}) = \sigma^2(A_d) + (n-1)^{1-d} [2\sigma(A_d, A_{s,2}) + (n-1)^{1-d}\sigma^2(A_{s,2})] \quad (18.10d)$$

so that (for sufficiently large n), the additive total variance increases provided $d < 1$. Both Hadfield and Wilson (2007) and Bijma (2010b) suggest methods to estimate the quantities in Equations 18.9 and 18.10a, respectively.

SELECTION UNDER THE PRESENCE OF ASSOCIATIVE EFFECTS

One of the key results when associative effects are present is that individual selection can result in a reversed response, while group selection always results in a positive response (although it may be far from optimal). These points were clearly made by Griffing (1967) for the simple case of two interacting, and unrelated, individuals within each group. For selection on individual phenotype, the response becomes

$$R = \frac{\bar{t}}{\sigma_z} [\sigma^2(A_d) + \sigma(A_d, A_s)] \quad (18.11a)$$

A negative covariance between direct and associative effects reduces the efficiency of selection, and if sufficiently negative, gives a reversed response. This loss of efficiency occurs because the only information an individual's phenotype contains about their breeding value for associative effects is that provided by the covariance between direct and associative breeding values (which can be negative). Conversely, if we select based on the mean of a group, we are selecting on both direct and associative effects to improve trait value. For the case of $n = 2$, Griffing obtained the expected response as

$$R = \frac{\bar{t}}{2\sigma(\bar{z})} [\sigma^2(A_d) + 2\sigma(A_d, A_s) + \sigma^2(A_s)] = \frac{\bar{t}}{2\sigma(\bar{z})} \sigma^2(A_T) \quad (18.11b)$$

While group selection always give a non-negative response, if the associative effects are weak, this approach is very inefficient relative to individual selection. For example, in the absence of associative effects, $\sigma^2(\bar{z}) = \sigma^2(z)/2$, and Equation 18.11b reduces to $\bar{t}\sigma(A_d)/[\sqrt{2}\sigma(z)]$, or $1/\sqrt{2} = 0.701$ of the response under individual selection.

Individual Selection: Theory

Consider individual selection in a group of size n that are potentially related. Recalling Equation 18.4a, the correlation between phenotype and total breeding value is

$$\sigma(z_i, A_{T_i}) = \sigma^2(A_d) + (n-1)\sigma(A_d, A_s) + \sum_{j \neq i} \sigma(A_{s_j}, A_{d_i} + (n-1)A_{s_i})$$

When individuals within the group are related, $\sigma(A_{s_j}, A_{s_i}) \neq 0$ and likewise if A_d and A_s are correlated, then for relatives we also have $\sigma(A_{d_j}, A_{s_i}) \neq 0$. If r_{ij} denotes the relationship between individuals i and j , then

$$\begin{aligned} \sum_{j \neq i} \sigma(A_{s_j}, A_{d_i} + (n-1)A_{s_i}) &= \sum_{j \neq i} \sigma(A_{s_j}, A_{d_i}) + (n-1) \sum_{j \neq i} \sigma(A_{s_j}, A_{s_i}) \\ &= \sigma(A_d, A_s) \sum_{j \neq i} r_{ij} + (n-1)\sigma^2(A_s) \sum_{j \neq i} r_{ij} \\ &= [\sigma(A_d, A_s) + (n-1)\sigma^2(A_s)] \left(\sum_{j \neq i} r_{ij} \right) \end{aligned} \quad (18.12a)$$

When all of the group members have the same relatedness ($r_{ij} = r$), the sum becomes $(n-1)r$, giving the result of Bijma et al. (2007a)

$$\sigma(z, A_T) = \sigma^2(A_d) + (n-1) [\sigma(A_d, A_s) + r\sigma(A_s, A_d) + r(n-1)\sigma^2(A_s)] \quad (18.12b)$$

$$= \sigma^2(A_d) + (n-1)(1+r)\sigma(A_d, A_s) + r(n-1)^2\sigma^2(A_s) \quad (18.12c)$$

Equation 18.12c shows the impact of having relatives within the group, which is to shift some of the variance in social effects $\sigma^2(A_s)$ into the covariance $\sigma(z, A_T)$ between individual phenotype and total breeding value. Use of relatives in the group thus allows individual selection to access some of this otherwise untapped variance. This shift arises because in a group of related individuals, their breeding values for social effects (which impact your phenotypic value) are now correlated with your own breeding values for social effects (which have no direct impact on your phenotype). A useful alternative expression is to break this covariance into that expected for unrelated groups (Equation 18.4b) plus the additional contribution due to individuals in the group being related,

$$\sigma(z, A_T) = \sigma(z, A_T | r = 0) + (n-1)r [\sigma(A_s, A_d) + (n-1)\sigma^2(A_s)] \quad (18.12d)$$

Alternatively, we can also express this as

$$\sigma(z, A_T) = r\sigma^2(A_T) + (1-r) [\sigma^2(A_d) + (n-1)\sigma(A_d, A_s)] \quad (18.12e)$$

The more closely related group members are, the more weight individual selection puts on A_T . In the extreme when groups are composed of clones, $\sigma(z, A_T) = \sigma^2(A_T)$. Plant breeding often selects among groups of clones, with such settings exploiting all of the heritable variation in both direct and associative effects without requiring any special designs.

Similarly, when all members in the group have the same relatedness r , the phenotypic variance becomes

$$\begin{aligned} \sigma^2(z) &= \sigma^2(A_d) + \sigma^2(E_d) + (n-1) [\sigma^2(A_s) + \sigma^2(E_s)] \\ &\quad + (n-1)r [2\sigma(A_s, A_d) + (n-2)\sigma^2(A_d)] \end{aligned} \quad (18.13a)$$

$$= \sigma^2(z | r = 0) + (n-1)r [2\sigma(A_s, A_d) + (n-2)\sigma^2(A_d)] \quad (18.13b)$$

where the phenotypic variance $\sigma^2(z | r = 0)$ when all group members are unrelated given by Equation 18.5c.

The response to selection is simply the change in the mean total breeding value, which (from Chapter 10) is the within-generation change in the phenotypic mean after selection (the selection differential S) times the slope of the regression of A_T on phenotype z ,

$$R = \frac{\sigma(z, A_T)}{\sigma_z^2} S = \frac{\sigma(z, A_T)}{\sigma_z} \bar{i} \quad (18.14)$$

with the second formulation following from the standard identity that $S = \sigma_z \bar{i}$ (Equation 10.6a). For $n = 2$ and $r = 0$, we recover Griffin's result (Equation 18.11a).

Example 18.4. Muir (2005) estimated variance components for six-week body weight in Japanese quail (*Coturnix coturnix japonica*) housed in groups of $n = 16$ per cage. REML estimates of the genetic variances were $\sigma^2(A_d) = 33.7$ and $\sigma^2(A_s) = 2.87$, while $\sigma(A_d, A_s) = -5.5$. Under these values, the predicted response to individual selection in a group of 16 unrelated individuals is

$$R = \frac{\bar{i}}{\sigma_z} [\sigma^2(A_d) + (n-1)\sigma(A_d, A_s)] = \frac{\bar{i}}{\sigma_z} [33.7 + 15 \cdot (-5.5)] = -48.8 \frac{\bar{i}}{\sigma_z}$$

The strong negative covariance between direct and social (competitive) effects results in an expected reversed response if directional selection is used, as the positive gain from improvement of direct effects is swamped by the negative effects from the correlated response in social values.

The presence of relatives within the group results in some fraction of $\sigma^2(A_s)$ being incorporated into the response under individual selection. Suppose the group of 16 consists of two half-sib families. In this case, the average relationship is 0.125, and from Equation 18.12d the resulting covariance between phenotype and total breeding values becomes

$$\begin{aligned}\sigma(z, A_T) &= \sigma(z, A_T | r = 0) + (n - 1)r [\sigma(A_s, A_d) + (n - 1)\sigma^2(A_s)] \\ &= -48.4 + 15 \cdot 0.125 (-5.5 + 15 \cdot 2.87) = 21.6\end{aligned}$$

Simply by using groups of relatives (as opposed to groups of unrelated individuals) allows individual selection to give an expected positive response.

Example 18.5. Consider a trait with $\sigma^2(A_d) = 500$, $\sigma^2(A_s) = 50$, $r(A_d, A_s) = -0.25$, and $\sigma^2(E_d) = 400$. For ease of presentation, we assume no social environmental effects (no E_s) and that E_d is uncorrelated across family members (i.e., no common family environment and no dominance). Consider a group size of $n = 6$. Given a selection intensity \bar{z} , what response is expected under individual selection when group members are unrelated? Here $\sigma(A_d, A_s) = -0.25 \sqrt{500 \cdot 50} = -39.5$. Substituting into Equation 18.4b gives the covariance as

$$\sigma(z, A_T) = \sigma^2(A_d) + (n - 1)\sigma(A_d, A_s) = 500 - 5 \cdot 39.5 = 302.5.$$

Likewise, from Equation 18.5c, the phenotypic variance is

$$\begin{aligned}\sigma^2(z) &= \sigma^2(A_d) + \sigma^2(E_d) + (n - 1) [\sigma^2(A_s) + \sigma^2(E_s)] \\ &= 500 + 400 + 5 \cdot 50 = 1150.\end{aligned}$$

Applying Equation 18.14 gives the resulting response as

$$R = \frac{\sigma(z, A_T)}{\sigma_z} \bar{z} = \frac{302.5}{\sqrt{1150}} \bar{z} = 8.92 \bar{z}.$$

Now suppose that group members are half-sibs ($r = 0.25$). What is the expected response? Applying Equation 18.12d, the covariance becomes

$$\begin{aligned}\sigma(z, A_T) &= \sigma(z, A_T | r = 0) + (n - 1)r [\sigma(A_s, A_d) + (n - 1)\sigma^2(A_s)] \\ &= 302.5 + 5 \cdot 0.25 (-39.5 + 5 \cdot 50) = 565.5,\end{aligned}$$

while Equation 18.13b gives the phenotypic variance as

$$\begin{aligned}\sigma^2(z) &= \sigma^2(z | r = 0) + (n - 1)r [\sigma(A_s, A_d) + (n - 2)\sigma^2(A_d)] \\ &= 1150 + 5 \cdot 0.25 (-39.5 + 4 \cdot 50) = 1350.6,\end{aligned}$$

yielding a response of $R = 15.39 \bar{z}$. Likewise, if the group consists of full sibs ($r = 0.5$), the resulting covariance, variance, and response are, respectively, 828.5, 1551.2, and $21.04 \bar{z}$. Response to selection increases with the relatedness of group members, with a 1.7- and 2.4-fold increase using groups of half- and full-sibs (respectively) relative to groups of unrelated individuals.

Individual Selection: Direct vs. Social Response

Recalling Equation 18.2b, the response in the trait has two components, the response in direct effects $R_d = \Delta\mu_{A_d}$ and the response in the social effects $R_s = \Delta\mu_{A_s}$. The relative contribution of each to the total response easily follows by considering the covariance of an individual's phenotype value z with either its direct A_d , or social A_s , breeding values. Specifically,

$$R_z = R_d + (n-1)R_s, \quad \text{where} \quad R_d = \frac{\sigma(A_d, z)}{\sigma_z} \bar{z} \quad \text{and} \quad R_s = \frac{\sigma(A_s, z)}{\sigma_z} \bar{z} \quad (18.15a)$$

Here

$$\sigma(A_d, z) = \sigma\left(A_d, A_d + \sum_{i \neq j} A_{s,i} + e\right) = \sigma^2(A_d) + r(n-1)\sigma(A_d, A_s) \quad (18.15b)$$

while

$$\sigma(A_s, z) = \sigma\left(A_s, A_d + \sum_{i \neq j} A_{s,i} + e\right) = \sigma(A_d, A_s) + r(n-1)\sigma^2(A_s) \quad (18.15c)$$

Equation 18.15b shows that the group must contain relatives ($r \neq 0$) in order for the covariance between direct and social values to impact the response in the direct value. Likewise, under individual selection, response in the social value only occurs if the direct and social values are correlated within individuals ($\sigma(A_d, A_s) \neq 0$) and/or if group members are related, in which case the social value of the focal individual is correlated with the social values of those within its group.

Example 18.6. Consider the response in a family of half-sibs from Example 18.5, where the expected total response was $15.39 \bar{z}$. What were the contributions from the direct and social response? For the values used in that example,

$$\sigma(A_d, z) = \sigma^2(A_d) + r(n-1)\sigma(A_d, A_s) = 500 + 0.25 \cdot 5 \cdot (-39.5) = 450.63$$

and

$$\sigma(A_s, z) = \sigma(A_d, A_s) + r(n-1)\sigma^2(A_s) = -39.5 + 0.24 \cdot 5 \cdot 50 = 23.0$$

Recalling (for half-sibs) that $\sigma_z^2 = 1350.6$, Equation 18.15 gives the two components of response as

$$R_d = \frac{450.63}{\sqrt{1350.6}} \bar{z} = 12.26 \bar{z}, \quad \text{and} \quad R_s = \frac{23}{\sqrt{1350.6}} \bar{z} = 0.63 \bar{z}$$

Hence, 80% ($12.26/15.39$) of the total response was due to response in direct effects, while 20% was from the response in social effects ($5 \cdot 0.63/15.39$). Under individual selection on half-sib families, both the mean direct and mean social values improved. By contrast, if group members are unrelated, then (Example 18.5) $\sigma_z^2 = 1150$, while

$$\sigma(A_d, z) = \sigma^2(A_d) = 500, \quad \sigma(A_s, z) = \sigma(A_d, A_s) = -39.5$$

giving responses of

$$R_d = \frac{500}{\sqrt{1150}} \bar{z} = 14.744 \bar{z}, \quad \text{and} \quad R_s = \frac{-39.5}{\sqrt{1150}} \bar{z} = -1.165 \bar{z}$$

While the total response in this case was positive, the large direct response (14.74) was significantly offset by a decrease in the mean social environment ($5 \cdot [-1.16] = -5.83$), giving

the total response as $(14.74 - 5.82) \bar{i} = 8.92 \bar{i}$. The lack of relatedness implies no direct selection involving $\sigma^2(A_s)$, and hence the social breeding values only change through their correlation with the direct values, which in this example was negative.

Individual Selection: Maternal Effects

An important special case, indeed the forerunner of more general models of associative effects, is the model of direct and **maternal** effects (Dickerson 1947; Willham 1963, 1972; Cheverud 1984). Here, the trait value of an individual is the function of its direct effect P_d and a **maternal performance trait** P_m contributed by its mother, so that if j is the mother of i , then

$$z_i = P_{d_i} + P_{m,j} \quad (18.16a)$$

In the absence of inbreeding, $r = 1/2$ for this group (mother-offspring) with $n = 2$. From Equation 18.12c, the covariance between phenotype and total breeding value ($A_T = A_d + A_m$),

$$\sigma(z, A_T) = \sigma^2(A_d) + (3/2)\sigma(A_d, A_m) + (1/2)\sigma^2(A_m), \quad (18.16b)$$

while Equation 18.13a gives the phenotypic variance as

$$\sigma^2(z) = \sigma^2(A_d) + \sigma(A_d, A_s) + \sigma^2(A_s) + \sigma_e^2 \quad (18.16c)$$

giving the resulting response to selection as

$$R = \frac{\sigma(z, A_T)}{\sigma_z} \bar{i} = \frac{\sigma^2(A_d) + (3/2)\sigma(A_d, A_m) + (1/2)\sigma^2(A_m)}{\sqrt{\sigma^2(A_d) + \sigma(A_d, A_s) + \sigma^2(A_s) + \sigma_e^2}} \bar{i} \quad (18.16d)$$

The total response can also be expressed in terms of the direct and maternal effect response. From Equation 18.15,

$$R_d = \frac{\sigma(A_d, z)}{\sigma_z} \bar{i} = \frac{\sigma^2(A_d) + (1/2)\sigma(A_d, A_m)}{\sigma_z} \bar{i} \quad (18.17a)$$

and

$$R_m = \frac{\sigma(A_m, z)}{\sigma_z} \bar{i} = \frac{\sigma(A_d, A_m) + (1/2)\sigma^2(A_m)}{\sigma_z} \bar{i} \quad (18.17b)$$

As reviewed by Cheverud (1984), most estimates of the direct-maternal covariance are negative. This raises the possibility of a reversed response due to a greater reduction in the maternal environment than improvement in the direct effect. It also allows for the trait to improve (via its direct value) at the expense of a declining maternal values.

We have already seen (Chapter 11) trait-based models of maternal effects, which have much more complicated dynamics (such as time lags). Chapter 37 examines these in more detail.

Group Selection: Theory

Under individual selection with unrelated group members, there is no contribution from $\sigma^2(A_s)$ to the response, and changes in A_s only enter as a correlated response to changes in A_d , which can be in an unfavorable direction. As we will see, $\sigma^2(A_s)$ enters into response under group selection even when there are no relatives in the group. The reason is that the group phenotype is a function of the distribution of A_s values.

Under strict group selection, selection is based on the group mean \bar{z} , or equivalently the total value of the group $n\bar{z} = \sum z$, and we will usually work with the latter. To obtain the covariance between the total value of a group and the total breeding value of one of its members, first note that

$$\begin{aligned} \sum_{j=1}^n z_j &= \sum_{j=1}^n \left(A_{d_j} + E_{d_j} + \sum_{k \neq j}^n (A_{s_k} + E_{s_k}) \right) = \sum_{j=1}^n A_{d_j} + \sum_{j=1}^n \sum_{k \neq j}^n A_{s_k} + \sum_{j=1}^n e_j \\ &= \sum_{j=1}^n A_{d_j} + (n-1) \sum_{j=1}^n A_{s_j} + \sum_{j=1}^n e_j = \sum_{j=1}^n A_{T_j} + \sum_{j=1}^n e_j \end{aligned} \quad (18.18)$$

where the residual values e_i sweep up a variety of environmental terms, and are given by Equation 18.1c. The residuals are assumed uncorrelated with any breeding values, but of course these can (and usually are) correlated within a group (e.g., Equation 18.23a). If r_{ij} is the relationship between individuals i and j , the covariance between the group total and the total breeding value of a group member is

$$\begin{aligned} \sigma \left(A_{T_i}, \sum_{j=1}^n z_j \right) &= \sigma \left(A_{T_i}, \sum_{j=1}^n (A_{T_j} + e_j) \right) = \sum_{j=1}^n \sigma(A_{T_i}, A_{T_j}) = \sigma^2(A_T) \sum_{j=1}^n r_{ij} \\ &= \sigma^2(A_T) \left(1 + \sum_{j \neq i} r_{ij} \right) \end{aligned} \quad (18.19a)$$

If group members are unrelated, then

$$\sigma \left(A_{T_i}, \sum_{j=1}^n z_j \right) = \sigma^2(A_T) \quad (18.19b)$$

which implies $\sigma(A_{T_i}, \bar{z}) = \sigma^2(A_T)/n$. Hence, group selection acts on the total breeding value of an individual, rather than on only part of this as is the case with individual selection (e.g., Equation 18.12e). The associative effects contribution to the total breeding value does not influence the phenotype of the focal individual, but does influence the phenotype of other group members. Group selection directly targets these effects. If all members have the same degree of relationship r ,

$$\sigma \left(A_{T_i}, \sum_{j=1}^n z_j \right) = \sigma^2(A_T) [1 + (n-1)r] \quad (18.19c)$$

Comparison of 18.19b and c shows the role of individuals within the group being related. Selection can proceed on associative effects even if none of the individuals in the group are related, but its efficiency is amplified by using relatives. From Equation 18.19c, the covariance with respect to the group mean becomes

$$\sigma(A_{T_i}, \bar{z}) = \frac{1}{n} \sigma^2(A_T) [1 + (n-1)r] = \sigma^2(A_T) \left(r + \frac{1-r}{n} \right) \quad (18.19d)$$

Turning now to the variance of the group total, a little bit of algebra is required. From Equation 18.18, we can decompose this group variance into additive-genetic and environmental components,

$$\sigma^2 \left(\sum_{j=1}^n A_{T_j} + \sum_{j=1}^n e_j \right) = \sigma^2 \left(\sum_{j=1}^n A_{T_j}, \sum_{k=1}^n A_{T_k} \right) + \sigma^2 \left(\sum_{j=1}^n e_j, \sum_{k=1}^n e_k \right) \quad (18.20)$$

Tackling the genetic component first,

$$\sigma\left(\sum_{j=1}^n A_{T_j}, \sum_{k=1}^n A_{T_k}\right) = \sigma^2(A_T) \cdot \sum_{j=1}^n \sum_{r=1}^n r_{ij} \quad (18.21a)$$

When all group members had the same relationship r , this reduces to

$$\sigma\left(\sum_{j=1}^n A_{T_j}, \sum_{k=1}^n A_{T_k}\right) = \sigma^2(A_T) n [1 + (n-1)r] \quad (18.21b)$$

Turning our attention to the residual terms, recall (Equation 18.1c) that the residual is a function of both direct and social environmental effects,

$$e_i = E_{d_i} + \sum_{k \neq i} E_{s_k}$$

Clearly, individuals within the same group are correlated because they share the E_s values from the other group members. First, the residual variance is

$$\begin{aligned} \sigma_e^2 = \sigma(e_i, e_i) &= \sigma\left(E_{d_i} + \sum_{k \neq i} E_{s_k}, E_{d_i} + \sum_{k \neq i} E_{s_k}\right) = \sigma(E_{d_i}, E_{d_i}) + \sum_{k \neq i} \sigma(E_{s_k}, E_{s_k}) \\ &= \sigma^2(E_d) + (n-1)\sigma^2(E_s) \end{aligned} \quad (18.22a)$$

For $i \neq j$ in the same group,

$$\begin{aligned} \sigma(e_i, e_j) &= \sigma\left(E_{d_i} + E_{s_j} + \sum_{k \neq i, j} E_{s_k}, E_{d_j} + E_{s_i} + \sum_{k \neq i, j} E_{s_k}\right) \\ &= \sigma(E_{d_i}, E_{d_j}) + \sigma(E_{d_i}, E_{s_i}) + \sigma(E_{d_j}, E_{s_j}) + \sigma\left(\sum_{k \neq i, j} E_{s_k}, \sum_{k \neq i, j} E_{s_k}\right) \\ &= 0 + 2\sigma(E_d, E_s) + \sum_{k \neq i, j} \sigma(E_{s_k}, E_{s_k}) \\ &= 2\sigma(E_d, E_s) + (n-2)\sigma^2(E_s) \end{aligned} \quad (18.22b)$$

The first term accounts for the fact that the direct and social environmental values can be correlated within the same individual, while the second term accounts for the shared environmental values contributed by the other $n-2$ group members. Putting these together gives

$$\sigma(e_i, e_j) = \begin{cases} \sigma_e^2 & i = j \\ \rho \sigma_e^2 & i \neq j \text{ for } i \text{ and } j \text{ in the same group} \end{cases} \quad (18.23a)$$

with

$$\sigma_e^2 = \sigma^2(E_d) + (n-1)\sigma^2(E_s), \quad \text{and} \quad \rho = \frac{2\sigma(E_d, E_s) + (n-2)\sigma^2(E_s)}{\sigma_e^2} \quad (18.23b)$$

Here ρ is the correlation among environmental values within a group, and can be either positive or negative. For large n , we expect the variance term to dominate the covariance term, giving $\rho > 0$. Equations 18.23a and b were first obtained by Bijma et al (2007b). Correlations among environmental residuals are also generated by shared maternal effects

and (for full-sibs) dominance. If all group members are the same type of relative, this is just incorporated into ρ . However, when a group consists of two (or more) families, the additional residual covariance among sibs need to be accounted for (see below for an example of this in a BLUP framework).

Using these results and following the same logic as with additive values gives

$$\sigma^2\left(\sum_{j=1}^n e_j, \sum_{k=1}^n e_k\right) = n\sigma_e^2 + \sum_{j \neq k} \sigma(e_j, e_k) = n\sigma_e^2 [1 + (n-1)\rho] \quad (18.24)$$

Substituting Equations 18.21b and 18.24 into Equation 18.20 gives the variance of the group total as

$$\sigma^2\left(\sum_{j=1}^n z_j\right) = n\sigma^2(A_T) [1 + (n-1)r] + n\sigma_e^2 [1 + (n-1)\rho] \quad (18.25a)$$

The variance of the group mean is just $1/n^2$ of this value, or

$$\begin{aligned} \sigma^2(\bar{z}) &= \sigma^2(A_T) \left(\frac{1 + (n-1)r}{n}\right) + \sigma_e^2 \left(\frac{1 + (n-1)\rho}{n}\right) \\ &= \sigma^2(A_T) \left(r + \frac{1-r}{n}\right) + \sigma_e^2 \left(\rho + \frac{1-\rho}{n}\right) \end{aligned} \quad (18.25b)$$

Using the covariance between total breeding value and group mean (Equation 18.19d) and the variance of the group mean (Equation 18.25b), the resulting response to selection becomes

$$R = \frac{\sigma(A_T, \bar{z})}{\sigma^2(\bar{z})} S = \frac{\sigma^2(A_T)r_n}{\sigma^2(A_T)r_n + \sigma_e^2\rho_n} S \quad (18.26a)$$

$$= \frac{\sigma(A_T, \bar{z})}{\sigma(\bar{z})} \bar{i} = \frac{\sigma^2(A_T)r_n}{\sqrt{\sigma^2(A_T)r_n + \sigma_e^2\rho_n}} \bar{i} \quad (18.26b)$$

where

$$r_n = r + \frac{1-r}{n} \quad \text{and} \quad \rho_n = \rho + \frac{1-\rho}{n}$$

For $n = 2$ and $r = 0$, applying Equations 18.19b and 18.25a, recovers Griffing's result (Equation 18.11b). As expected, with only direct effects, these reduce to our expressions for family selection (Chapter 17).

Example 18.7. Consider group selection using Muir's quail data from Example 18.4. Here $\sigma^2(A_d) = 33.7$, $\sigma^2(A_s) = 2.87$, $\sigma(A_d, A_s) = -5.5$, $n = 16$. Muir estimated the residual variance as $\sigma_e^2 = 69.0$, while Muir's model assumed $\rho = 0$, giving $\rho_n = 1/n$ and hence $\sigma_e^2\rho_n = 69.0/16 = 4.32$. Applying Equation 18.4 gives the total additive variance as

$$\begin{aligned} \sigma^2(A_T) &= \sigma^2(A_d) + 2(n-1)\sigma(A_d, A_s) + (n-1)^2\sigma^2(A_s) \\ &= 33.7 + 30 \cdot (-5.5) + 30^2 \cdot 2.87 = 2451.7, \end{aligned}$$

while Equation 18.26b gives the response as

$$R = \frac{\sigma^2(A_T)r_n}{\sqrt{\sigma^2(A_T)r_n + \sigma_e^2\rho_n}} \bar{i} = \frac{2451.7 \cdot r_n}{\sqrt{2451.7 \cdot r_n + 4.32}} \bar{i}$$

For groups of unrelated individuals, $r = 0$ and $r_n = 0.0625$, and the response becomes $R = 12.2\bar{i}$. For half- and full-sibs, $r_n = 0.297$ and 0.531 , with responses of $26.9\bar{i}$ and $36.0\bar{i}$, a two- and three-fold increase relative to a group of unrelated individuals.

While Equation 18.26 shows that group selection always results in an expected non-negative response (as $\sigma^2(A_T) \geq 0$), it may be less than optimal. If direct effects account for the majority of variance, group selection can be very inefficient relative to individual selection. To see this, consider groups of unrelated individuals and suppose the trait of interest has no associative effects [$\sigma^2(A_s) = 0$], so that $\sigma^2(A_T) = \sigma^2(A_d)$. Under individual selection the response is $R_I = h\sigma(A_d)\bar{i}$ (Equation 10.6b). Now consider the response under group selection, where $\sigma(\bar{z}, A_T) = (1/n)\sigma^2(A_d)$ and $\sigma^2(\bar{z}) = \sigma_z^2/n$, giving the response as

$$R_G = \frac{\sigma(\bar{z}, A_T)}{\sigma^2(\bar{z})} \bar{i} = \frac{(1/n)\sigma^2(A_d)}{\sigma_z^2/\sqrt{n}} \bar{i} = \frac{1}{\sqrt{n}} \frac{\sigma(A_d)}{\sigma_z} \sigma(A_d) \bar{i} = \frac{1}{\sqrt{n}} h\sigma(A_d) \bar{i} = \frac{1}{\sqrt{n}} R_I$$

Under these conditions, individual selection is always superior to group selection, with the superiority increasing with group size. For groups of 5, 10, and 25, group selection has only 44.7%, 31.6%, and 20% (respectively) of the expected response of individual selection.

Group Selection: Direct vs. Social Response

As was the case for individual selection, we can decompose the response under group selection into the responses from direct and social effects, $R_z = R_d + (n-1)R_s$. Under group selection, these response components are given by

$$R_d = \frac{\sigma(A_d, \sum z_j)}{\sigma(\sum z_j)} \bar{i} \quad \text{and} \quad R_s = \frac{\sigma(A_s, \sum z_j)}{\sigma(\sum z_j)} \bar{i} \quad (18.27a)$$

The covariance between the direct breeding value and group total becomes

$$\begin{aligned} \sigma\left(A_{d_i}, \sum_{j=1}^n z_j\right) &= \sigma\left(A_{d_i}, \sum_{j=1}^n A_{d_j} + (n-1) \sum_{j=1}^n A_{s_j} + \sum_{j=1}^n e_j\right) \\ &= \sigma^2(A_d) \sum_{j=1}^n r_{ij} + (n-1)\sigma(A_d, A_s) \sum_{j=1}^n r_{ij} \\ &= [\sigma^2(A_d) + (n-1)\sigma(A_d, A_s)] [1 + (n-1)r] \end{aligned} \quad (18.27b)$$

Similarly, for the social breeding value

$$\sigma\left(A_{s_i}, \sum_{j=1}^n z_j\right) = [\sigma(A_d, A_s) + (n-1)\sigma^2(A_s)] [1 + (n-1)r] \quad (18.27c)$$

Increasing the relatedness r of group members increases the response in both components by the same proportional amount, $[1 + (n-1)r]$. Hence, the relative contribution of these two components is independent of the degree of relatedness within the group. By contrast, recall (Equations 18.15b,c) that under individual selection, the relative contributions of these two components changes (and potentially rather dramatically) with r .

Group Selection: Experiment Evidence

How effective is group selection? As reviewed in Chapter 17, the special case of the group being a single family has a fairly robust experimental literature. What is seen in more general settings? Experiments in laboratory settings generally have proved effective in generating a positive response (Goodnight and Stevens 1997, Goodnight 2005). Especially telling are several reports of group selection giving a positive response when individual selection either failed or generated a negative response.

One of the first experiments was by Wade (1976, 1977), who found a rapid response to group selection for population size in the flour beetle *Tribolium castaneum*. A series of 48 populations were founded, each with 16 unrelated individuals, and population size was measured at 37 days post-founding. Under the control (allowing for individual selection during the grow-out to 37 days), a group of 16 individuals was chosen at random from the 48 populations and used to found a new population, repeating (with the possibility of resampling from the same population) until all new 48 populations are formed. Under group selection for increased population size, sets of 16 individuals are drawn from the largest population and used to found a new population, which is continued until the largest population is exhausted. When this happens, individuals are similarly used from the second largest population, and so forth to fill out the new array of 48 populations. The group-selected populations showed a significantly larger sizes relative to the control, and also showed reduced levels of cannibalism. Laboratory populations of *Tribolium* were also used by Craig (1982), who also found that group selection was very efficient in increasing (and decreasing) emigration rates. In both studies, some degree of relationship might be expected within groups, which would be small at first, but increasing under inbreeding as selection proceeds (albeit still somewhat small at the end of the experiment).

Response under group selection is not limited to animals. Goodnight (1985) contrasted individual vs. group selection for leaf area in the mustard *Arabidopsis thaliana*. Plants were grown in groups of 16 unrelated individuals. Individual selection for increased leaf area actually resulted in a reversed response, with offspring showing smaller leaf area. In contrast, average leaf area per plant (i.e., a larger total leaf area for the group) increased under group selection.

Finally, dramatic responses with significant economic impact have occurred using group selection in production settings. In chickens, high egg production systems typically house several hens per cage. However, aggressive behavior and mortality is common in such settings. Selection for improved individual production could result in increased aggression within the cage, and hence lower production (reviewed in Muir 1985). To see if group selection could indeed improve performance, Muir (1996) selected based on cages of nine birds, where all birds within a given cage had the same size. Eggs per hen per day, eggs per hen, and egg mass all increased dramatically. More striking, annual percentage mortality declined from 68% to just under 9% at the end of generation 6, which is similar to the mortality in single-bird cages. Muir called the resulting selected strain KGB chickens (for Kinder, Gentler Birds). Selection based on the group improved total performance of the cage, in part by reducing the amount of aggression within the cage. Using the KGB line, Craig and Muir (1996) found that they were indeed less aggressive (beak-inflicted injuries were significantly reduced).

The benefits of group selection have often be framed in terms of exploiting non-additive variation that is not accessible to traditional individual selection (e.g., Goodnight and Stevens 1997). While we have focused here on genetic variation that is not directly accessible under individual selection when groups are unrelated ($\sigma^2[A_s]$), this variation is entirely *additive*. Specifically, when heritable associative effects are present, they can only be directly accessible through either group selection (with either related or unrelated members) or individual selection when interactions occur in groups of related individuals. This is not to ignore the possibility of exploiting additional non-additive variation under group selection, but rather to highlight the importance of associative effects. Thus, the focus for group selection shifts

from additional non-additive variation in a trait to additive genetic variation in a second trait (the associative effect) that can be exploited under appropriate selection designs (such as individual selection in groups of relatives or group selection).

INCORPORATING BOTH INDIVIDUAL AND GROUP INFORMATION

Given that group selection always results in an expected positive response, while individual selection can range from being far more efficient than group selection to generating an expected reversed response, clearly the optimal approach is some combination of selection on both individual and group components. This is simply an extension of the family index introduced in Chapter 17 that combines both individual and family information.

Response on a Weighted Index

To combine both individual and group selection, consider the index I , where the value of the index for the i th individual is

$$I_i = z_i + g \sum_{j \neq i} z_j \quad (18.27a)$$

This index can also be written as

$$I_i = (1 - g)z_i + g \sum_{j=1}^n z_j = (1 - g)z_i + g n \bar{z}, \quad (18.27b)$$

showing that individual selection corresponds to $g = 0$, while $g = 1$ corresponds to group selection. Thus, Equation 18.27b contains our previous results for individual and group selection as special cases. Also note that selection of individuals based on within-group deviations are also a special case of Equation 18.27a, by setting $g = -1/n$, giving

$$I_i = z_i - \frac{1}{n} \sum_{j=1}^n z_j = z_i - \bar{z} \quad (18.27c)$$

The response in μ_z to selection on this index is

$$R = \frac{\sigma(I, A_T)}{\sigma(I)} \bar{i}_I \quad (18.28a)$$

This can also be written in terms of the **accuracy** of selection, a concept first introduced in Chapter 10 (Equation 10.11), which is the correlation between the target of selection (here I) and the breeding objective (here the total breeding value A_T). We can express the ratio in Equation 18.28a as

$$\frac{\sigma(I, A_T)}{\sigma(I)} = \frac{\sigma(I, A_T)}{\sigma(I)} \frac{\sigma(A_T)}{\sigma(A_T)} = \sigma(A_T) \frac{\sigma(I, A_T)}{\sigma(A_T) \sigma(I)} = \sigma(A_T) \rho(A_T, I) \quad (18.28b)$$

where the accuracy $\rho(A_T, I)$ is the correlation between the index value of an individual and their total breeding value. Using this result, Equation 18.28a becomes

$$R = \rho(A_T, I) \sigma(A_T) \bar{i}_I \quad (18.28c)$$

This is a very useful expression for comparing different selection schemes as $\sigma(A_T)$ remains unchanged (provided group size remains fixed), so that the maximal response occurs by maximizing $\rho(A_T, I) \bar{t}_I$. Given that the fraction saved largely sets the selection intensity \bar{t}_I (subject to minor variation due to finite populations, see Equation 10.27b), the optimal scheme is that which maximizes the accuracy $\rho(A_T, I)$.

To obtain the general expression for response for any combination of group selection fraction g and relatedness within groups r , we first need the covariance of I and A_T within an individual. This is obtained as follows:

$$\sigma(A_T, I) = (1 - g)\sigma(A_T, z) + g\sigma\left(A_T, \sum_{j=1}^n z_j\right) \quad (18.29a)$$

When group members are unrelated, Equations 18.4b and 18.19b give

$$\sigma(A_T, I) = (1 - g) [\sigma^2(A_d) + (n - 1)\sigma(A_d, A_s)] + g\sigma^2(A_T) \quad (18.29a)$$

When group members all have the same relationship, Equations 18.12e and 18.19c give

$$\begin{aligned} \sigma(A_T, I) = (1 - g) & \left(r\sigma^2(A_T) + (1 - r) [\sigma^2(A_d) + (n - 1)\sigma(A_d, A_s)] \right) \\ & + g[1 + (n - 1)r] \sigma^2(A_T) \end{aligned} \quad (18.29b)$$

Collecting terms, this reduces to

$$\sigma(A_T, I) = [g + r + (n - 2)gr] \sigma^2(A_T) + (1 - g)(1 - r) [\sigma^2(A_s) + (n - 1)\sigma(A_s, A_d)] \quad (18.29c)$$

While parts of this result (in a bit more cryptic form) appear in Griffing (1977), the fully general version given by Equation 18.29c is due to Bijma et al. (2007a). Note that g and r have symmetric roles in the covariance between the index and total breeding value. Thus, from the standpoint of this covariance, relatedness and group selection (r and g) are interchangeable. However, as we will next demonstrate, g and r do not play symmetric roles in the variance σ_I^2 of the index, so that interchanging r and g values results in a different variance and hence a different response (Example 18.8).

Now consider the variance of the index I . From Equation 18.27a,

$$\begin{aligned} \sigma_I^2 &= \sigma\left(z_i + g \sum_{j \neq i} z_j, z_i + g \sum_{j \neq i} z_j\right) \\ &= \sigma_z^2 + 2g\sigma\left(z_i, \sum_{j \neq i} z_j\right) + g^2\sigma^2\left(\sum_{j \neq i} z_j\right) \end{aligned} \quad (18.30a)$$

If all group members have the same relationship, then

$$\sigma\left(z_i, \sum_{j \neq i} z_j\right) = (n - 1)\sigma(z_i, z_j) \quad (18.30b)$$

and

$$\sigma^2\left(\sum_{j \neq i} z_j\right) = (n - 1) [\sigma_z^2 + (n - 2)\sigma(z_i, z_j)] \quad (18.30c)$$

Substituting these two expressions into Equation 18.30a and collecting terms gives

$$\sigma_I^2 = \sigma_z^2 [1 + g^2(n - 1)] + \sigma(z_i, z_j)g(n - 1) [2 + g(n - 2)] \quad (18.30d)$$

As a check of Equation 18.30d, note that (as expected) this reduces to σ_z^2 for $g = 0$ and to $n\sigma_z^2 + n(n-1)\sigma(z_i, z_j)$ for $g = 1$. Equation 18.13a gives the expression for σ_z^2 when all relatives within the group are related by r . It remains to obtain $\sigma(z_i, z_j)$, the phenotypic covariance of group members, in order to apply Equation 18.30d. From Equation 18.1c,

$$\sigma(z_i, z_j) = \sigma\left(A_{d_i} + \sum_{k \neq i} A_{s_k} + e_i, A_{d_j} + \sum_{k \neq j} A_{s_k} + e_j\right) \quad (18.31a)$$

$$= \sigma(A_{d_i}, A_{d_j}) + 2\sigma\left(A_{d_i}, \sum_{k \neq i} A_{s_k}\right) + \sigma\left(\sum_{k \neq i} A_{s_k}, \sum_{k \neq j} A_{s_k}\right) + \sigma(e_i, e_j) \quad (18.31b)$$

Expanding and evaluating the above covariance terms and collecting common terms, this ultimately reduces to

$$\begin{aligned} \sigma(z_i, z_j) = & 2\sigma(A_d, A_s) + (n-2)\sigma^2(A_s) + \rho\sigma_e^2 \\ & + r[\sigma^2(A_d) + 2(n-2)\sigma(A_d, A_s) + \{(n-1) + (n-2)^2\}\sigma^2(A_s)] \end{aligned} \quad (18.31c)$$

Notice that the term appearing when group members are related ($r \neq 0$) is the variance of A_T when the group size is $(n-1)$ plus the correction $(n-1)\sigma^2(A_s)$. Equations 18.29 and 18.30 are substituted into Equation 18.28 to obtain the response. The interplay of $\sigma(A_T, I)$ and σ_I^2 (as functions of g and r) determine the accuracy of any particular index (Figure 18.3).

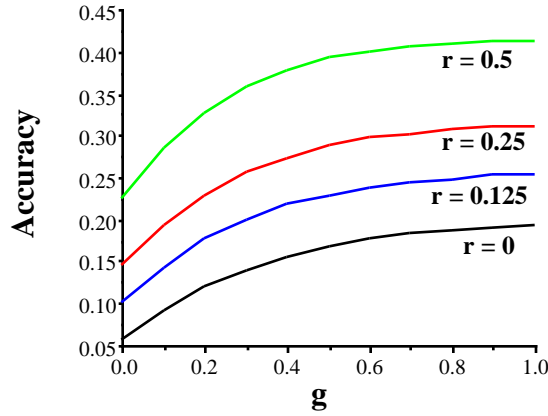


Figure 18.3. Accuracy of the index I as a function of the group weight g for groups of different types of relatives. The variance components used here are from Ellen et al. (2008) for survival days for caged chickens, and are given in Example 18.8, which also works through the calculations.

Example 18.8. Ellen et al. (2008) estimated the following values for survival days in chickens raised in groups of $n = 4$: $\sigma^2(A_d) = 915$, $\sigma(A_d, A_d) = 62$, $\sigma^2(A_s) = 134$, $\sigma^2(e) = 11500$, $\rho = 0.08$. Applying Equation 18.4c,

$$\begin{aligned} \sigma^2(A_T) &= \sigma^2(A_d) + 2(n-1)\sigma(A_d, A_s) + (n-1)^2\sigma^2(A_s) \\ &= 915 + 2 \cdot 3 \cdot 62 + 3^2 \cdot 134 = 2493, \end{aligned}$$

while

$$\sigma^2(A_s) + (n-1)\sigma(A_s, A_d) = 134 + 3 \cdot 62 = 320.$$

Substituting these results into Equation 18.29 gives the covariance between I and total breeding value of

$$\sigma(A_T, I) = (g + r + 2gr) \cdot 2493 + (1 - g)(1 - r) \cdot 320 \quad (18.32a)$$

To apply Equation 18.30d to obtain the variance in I , we first need expressions for $\sigma^2(z)$ and $\sigma(z_i, z_j)$. From Equation 18.5d, the phenotypic variance when the group contains unrelated individuals is

$$\sigma^2(z | r = 0) = \sigma^2(A_d) + (n-1)\sigma^2(A_s) + \sigma^2(e) = 915 + 3 \cdot 134 + 11,500 = 12,817$$

Since

$$2\sigma(A_s, A_d) + (n-2)\sigma^2(A_d) = 2 \cdot 62 + 2 \cdot 915 = 1954,$$

Equation 18.13b gives the phenotypic variance for a group with relationship r as,

$$\begin{aligned} \sigma^2(z) &= \sigma^2(z | r = 0) + (n-1)r [2\sigma(A_s, A_d) + (n-2)\sigma^2(A_d)] \\ &= 12,817 + r \cdot 5862 \end{aligned} \quad (18.32b)$$

To obtain $\sigma(z_i, z_j)$, we first find that

$$2\sigma(A_d, A_s) + (n-2)\sigma^2(A_s) + \rho\sigma_e^2 = 2 \cdot 62 + 2 \cdot 134 + 0.08 \cdot 11550 = 1312$$

and

$$\begin{aligned} \sigma^2(A_d) + 2(n-2)\sigma(A_d, A_s) + \{(n-1) + (n-2)^2\} \sigma^2(A_s) \\ = 915 + 2 \cdot 2 \cdot 62 + (3 + 2^2) \cdot 134 = 2101 \end{aligned}$$

Substituting into Equation 18.31c gives the general covariance between group members when all members are related by an amount r as

$$\sigma(z_i, z_j) = 1312 + r \cdot 2101 \quad (18.32c)$$

Finally, substituting Equations 18.32b,c into Equation 18.30d gives the variance in I as

$$\begin{aligned} \sigma_I^2 &= \sigma_z^2 [1 + g^2(n-1)] + \sigma(z_i, z_j)g(n-1) [2 + g(n-2)] \\ &= (12,817 + r \cdot 5862) (1 + 3g^2) + (1312 + r \cdot 2101) 6g (1 + g) \end{aligned} \quad (18.32d)$$

Note from 18.32a (and indeed 18.29b) that the roles of relatedness r and amount of group selection g are fully interchangeable in the covariance between I and A_T . However, Equation 18.32d shows that this is *not* the case for $\sigma^2(I)$, and hence the expected responses when the values of r and g are swapped are not expected to be equal. For example, for $r = 0.5, g = 0$,

$$\sigma(I, A_T) = 1406.5, \quad \sigma_I^2 = 15748, \quad \rho(A_T, I) = 0.224, \quad R = 11.21 \bar{i},$$

while for $r = 0, g = 0.5$,

$$\sigma(I, A_T) = 1406.5, \quad \sigma_I^2 = 28334, \quad \rho(A_T, I) = 0.167, \quad R = 8.36 \bar{i}.$$

As points of reference, the accuracy and response under individual selection with no relatives in the group ($g = r = 0$) are $\rho(A_T, I) = 0.057$ and $R = 2.83 \bar{i}$, while for group selection with unrelated individuals within the group ($g = 1, r = 0$), the accuracy and response become $\rho(A_T, I) = 0.193$ and $R = 9.63 \bar{i}$.

Optimal Response

In the index given by Equation 18.27a, g is the fraction of weight placed on a random individual from the group that interacts with the focal individual. If this weight is zero, the index reduces to individual selection, while if this weight is 1, all individuals in the group are weighted equally in the selection decision and we have group selection. As presented, Equation 18.27a implies a range on g from 0 to 1. An obvious question is the optimal value for g to maximize response. From Equation 18.28c, we see that the optimal response occurs by using those weights in I that maximizes the correlation $\rho(A_T, I)$ between I and A_T . To find these optimal weights, we start with the index

$$I = b_1 z + b_2 \sum_{j \neq i} z_j \quad (18.33a)$$

with no restrictions placed on the ranges of b_1 and b_2 . Selection on this index is equivalent to selection using the index

$$I = z + \frac{b_2}{b_1} \sum_{j \neq i} z_j \quad (18.33b)$$

Hence, the connection between Equations 18.27a and 18.33a is that $g = b_2/b_1$. The difference is that we no longer restrict consideration of g between zero and one. All of the previous results for selection response on Equation 18.27a hold for any value of g , but we focused on the zero-one range given the transition from individual to group selection. More generally, we could have negative weights, or a g value that exceeds one. In the former case, negative g values correspond to a weighted within-group deviation. In the latter case ($g > 1$), we place more weight on a random individual from the interacting group than the focal individual. This might occur when associative effects are much larger than direct effects, and hence group members provide more information than the focal individual about the value of A_T in that focal individual.

In Chapter 17, we were able to obtain straightforward expressions for optimal weights in a family index (Equation 17.54). Index selection theory (Chapter 33) gives the optimal index weights in the form of a matrix expression (Equations 18.35a, 33.18a), which greatly simplified under simple family selection (i.e., no associative effects). Unfortunately, such is not the case here, and so we (very briefly) introduce the machinery for obtaining an optimal index, with the theory fully developed in Chapters 33 and 34. The idea is that there are two potentially difference indices: the index I we selection on and the index H whose response we wish to maximize. Specifically, we select on some index $I = \mathbf{b}^T \mathbf{x}$ where x_i is the value of trait i used to make selection decisions and b_i the weight placed on that trait in the index. In keeping with Equation 18.33a, the vector of phenotypes for individual i is

$$\mathbf{x} = \begin{pmatrix} z_i \\ \sum_{j \neq i} z_j \end{pmatrix} \quad (18.34a)$$

Using this index to make selection decisions, we wish to find the weights \mathbf{b} to maximize response for some weighted combination of variables $H = \mathbf{a}^T \mathbf{g}$. Here the elements of \mathbf{a} are the weights and \mathbf{g} is the vector of breeding values for the traits of interest. In our case, we wish to maximize response in the total breeding value, so that

$$H = A_T = A_d + (n-1)A_s = \mathbf{a}^T \mathbf{g} \quad (18.34b)$$

where

$$\mathbf{g} = \begin{pmatrix} A_d \\ A_s \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} 1 \\ n-1 \end{pmatrix} \quad (18.34c)$$

The optimal weights \mathbf{b}_s to use in I to maximize response in H (i.e., obtain the highest correlation between A_T and I) are given by the Smith-Hazel index (Chapter 33, Equation 33.18a), where

$$\mathbf{b}_s = \mathbf{P}^{-1} \mathbf{G}^T \mathbf{a} \quad (18.35a)$$

\mathbf{P} is the phenotypic covariance matrix for the elements in \mathbf{x} , which in our case becomes

$$\mathbf{P} = \begin{pmatrix} \sigma^2(z) & \sigma(z_i, \sum_{j \neq i} z_j) \\ \sigma(z_i, \sum_{j \neq i} z_j) & \sigma(\sum_{j \neq i} z_j, \sum_{j \neq i} z_j) \end{pmatrix} \quad (18.36a)$$

$$= \begin{pmatrix} \sigma^2(z) & (n-1)\sigma(z_i, z_j) \\ (n-1)\sigma(z_i, z_j) & (n-1)[\sigma_z^2 + (n-2)\sigma(z_i, z_j)] \end{pmatrix} \quad (18.36b)$$

where we have used Equations 18.30b,c to obtain Equation 18.36b. \mathbf{G} is the matrix of covariances between traits in the index H and the index I , with $G_{ij} = \sigma(g_i, x_j)$. Since different traits are involved in the two indices, \mathbf{G} need not be symmetric. For our case,

$$\mathbf{G}^T = \begin{pmatrix} \sigma(z_i, A_{d_i}) & \sigma(z_i, A_{s_i}) \\ \sigma(\sum_{j \neq i} z_j, A_{d_i}) & \sigma(\sum_{j \neq i} z_j, A_{s_i}) \end{pmatrix} \quad (18.37a)$$

where

$$\sigma(z_i, A_{d_i}) = \sigma^2(A_d) + r(n-1)\sigma(A_d, A_s) \quad (18.37b)$$

$$\sigma(z_i, A_{s_i}) = \sigma(A_d, A_s) + r(n-1)\sigma^2(A_s) \quad (18.37c)$$

$$\sigma\left(\sum_{j \neq i} z_j, A_{d_i}\right) = (n-1)\sigma(A_d, A_s) + r(n-1)[\sigma^2(A_d) + (n-2)\sigma(A_d, A_s)] \quad (18.37d)$$

$$\sigma\left(\sum_{j \neq i} z_j, A_{s_i}\right) = (n-1)\sigma^2(A_s) + r(n-1)[\sigma(A_d, A_s) + (n-2)\sigma^2(A_s)] \quad (18.37e)$$

Equations 18.37b-e follow from the approach used throughout this chapter of a term-by-term evaluation of the covariance. The use of index selection machinery to solve this problem was initially outlined by Ellen et al. (2007).

Example 18.9. Consider a trait where $\sigma(A_d, A_s) = 0$, no correlation between environmental values within the group ($\rho = 0$) and no relatives in the group. Equation 18.5b gives $\sigma^2(z) = \sigma^2(A_d) + (n-1)\sigma^2(A_s) + \sigma^2(e)$, while Equation 18.31c gives $\sigma(z_i, z_j) = (n-2)\sigma^2(A_s)$. Hence,

$$\mathbf{P} = \begin{pmatrix} \sigma^2(z) & (n-1)(n-2)\sigma^2(A_s) \\ (n-1)(n-2)\sigma^2(A_s) & (n-1)[\sigma^2(z) + (n-2)^2\sigma^2(A_s)] \end{pmatrix}$$

Likewise, Equations 18.37c-e give

$$\mathbf{G} = \begin{pmatrix} \sigma^2(A_d) & 0 \\ 0 & (n-1)\sigma^2(A_s) \end{pmatrix}$$

What are the optimal weights for a trait with $\sigma^2(A_d) = 100$, $\sigma^2(A_s) = 9$, $\sigma^2(e) = 100$, $n = 10$? For these values, Equation 18.35a yields

$$\mathbf{b}_s = \mathbf{P}^{-1} \mathbf{G}^T \mathbf{a} = \begin{pmatrix} 281 & 648 \\ 648 & 9090 \end{pmatrix}^{-1} \begin{pmatrix} 100 & 0 \\ 0 & 81 \end{pmatrix} \begin{pmatrix} 1 \\ 9 \end{pmatrix} = \begin{pmatrix} 0.2046 \\ 0.0656 \end{pmatrix}$$

The resulting index weight g is $0.0656/0.2046 = 0.32$, giving the index as

$$I_i = z_i + 0.32 \cdot \sum_{j \neq i} z_j$$

If we increase $\sigma^2(A_s)$ to 15, we now have the optimal weights as

$$\mathbf{b}_s = \mathbf{P}^{-1} \mathbf{G}^T \mathbf{a} = \begin{pmatrix} 335 & 1080 \\ 1080 & 13950 \end{pmatrix}^{-1} \begin{pmatrix} 100 & 0 \\ 0 & 136 \end{pmatrix} \begin{pmatrix} 1 \\ 9 \end{pmatrix} = \begin{pmatrix} 0.0236 \\ 0.0852 \end{pmatrix}$$

giving a g value of $0.0852/0.0236 = 3.6$, and an optimal index of

$$I_i = z_i + 3.6 \cdot \sum_{j \neq i} z_j$$

Hence, on average, all of the individuals with which our focal individual interacts are given almost four times the weight as the phenotype of the focal individual. Finally, suppose $\sigma^2(A_s) = 20$. In this case

$$\mathbf{b}_s = \mathbf{P}^{-1} \mathbf{G}^T \mathbf{a} = \begin{pmatrix} 380 & 1440 \\ 1440 & 18000 \end{pmatrix}^{-1} \begin{pmatrix} 100 & 0 \\ 0 & 180 \end{pmatrix} \begin{pmatrix} 1 \\ 9 \end{pmatrix} = \begin{pmatrix} -0.112 \\ 0.0989 \end{pmatrix}$$

giving the optimal index as

$$I_i = 0.0989 \cdot \left(\sum_{j \neq i} z_j \right) - 0.112 \cdot z_i$$

Writing the summation as $n \bar{z} - 10 \bar{z}$, this is equivalent to selection on the index

$$I_i = 0.989 \bar{z} - 0.112 z_i$$

We can rewrite this index as

$$I_i = (0.989 - 0.112) \bar{z} - 0.112 (z_i - \bar{z}) = 0.877 \bar{z} - 0.112 (z_i - \bar{z})$$

which is equivalent to selecting using the index

$$I_i = \bar{z} - \frac{0.112}{0.877} (z_i - \bar{z}) = \bar{z} - 0.128 (z_i - \bar{z})$$

Hence, the optimal index in this case is the group mean minus a weighted within-group deviation.

While Griffing developed many of the basic equations for selection response, one reason for the low impact of this important work was that, at the time, there was no reliable way to estimate the key variance components $\sigma^2(A_d)$, $\sigma^2(A_s)$, and $\sigma(A_d, A_s)$. These are required to compare h_d^2 with T^2 , and hence to judge the potential of additional genetic variation that cannot be exploited under individual selection. Further, reasonable estimates of these variance components are required to obtain the optimal index weights. Finally, without some tangible values, Griffing's work was for some a bit too abstract – the observed phenotype written the sum of two unmeasured, and usually not estimated, components. The solution to these problems was suggested by Muir and Schinckel (2002), and detailed in the seminal paper of Muir (2005), who put these problems into a standard BLUP/REML mixed-model framework (Chapters 15, 16, 35; LW Chapters 26, 27).

Mixed-Model Estimation of Direct and Associative Effects

The general approach follows by considering a standard animal model with addition random effects (Equations 15.20-21). Equation 18.1b shows how the phenotype of individual i is the sum of its direct breeding value, the social breeding values of its group members, and the environmental effects,

$$z_i = \mu + (A_{d_i} + E_{d_i}) + \sum_{j \neq i} (A_{s_j} + E_{s_j})$$

To start with, we will assume a very simple residual structure,

$$z_i = \mu + A_{d_i} + \sum_{j \neq i} A_{s_j} + e_i$$

where the e_i are uncorrelated and homoscedastic. Letting \mathbf{a}_d denote the vector of **direct breeding values (DBVs)**, and \mathbf{a}_s the vector of **social breeding values (SBVs)**, the resulting mixed model becomes

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_d \mathbf{a}_d + \mathbf{Z}_s \mathbf{a}_s + \mathbf{e} \quad (18.38)$$

Here $\boldsymbol{\beta}$ is the vector of fixed effect (which is simply the mean for our simple example) and \mathbf{X} the design matrix associated with these fixed effects. Likewise, \mathbf{Z}_d and \mathbf{Z}_s are the corresponding incidence matrices for the direct and social effects, which follow easily by considering the group member (Examples 18.10, 18.11).

To complete the model, we need to specify the covariance structures of the three vectors of random effects. Our assumption on the residual errors implies that the covariance matrix for the residuals is $\sigma_e^2 \mathbf{I}$. The covariance structure for the two vectors of random effects is a function of the relationship matrix \mathbf{A} of the individuals in the study, with

$$V \begin{pmatrix} \mathbf{a}_d \\ \mathbf{a}_s \end{pmatrix} = \begin{pmatrix} \sigma^2(A_d)\mathbf{A} & \sigma(A_d, A_s)\mathbf{A} \\ \sigma(A_d, A_s)\mathbf{A} & \sigma^2(A_s)\mathbf{A} \end{pmatrix} \quad (18.39a)$$

This is often written more compactly using the **Kronecker** or **direct product** notation as $\mathbf{G} \otimes \mathbf{A}$ where

$$\mathbf{G} = \begin{pmatrix} \sigma^2(A_d) & \sigma(A_d, A_s) \\ \sigma(A_d, A_s) & \sigma^2(A_s) \end{pmatrix} \quad (18.39b)$$

Since the residuals are assumed uncorrelated with the other random effects, the full covariance structure for this model is

$$V \begin{pmatrix} \mathbf{a}_d \\ \mathbf{a}_s \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \sigma^2(A_d)\mathbf{A} & \sigma(A_d, A_s)\mathbf{A} & 0 \\ \sigma(A_d, A_s)\mathbf{A} & \sigma^2(A_s)\mathbf{A} & 0 \\ 0 & 0 & \sigma_e^2 \mathbf{I} \end{pmatrix} \quad (18.39c)$$

Example 18.10. To introduce how a mixed-model with direct and social effects is constructed, consider the following simple example where eight individuals are measured. Individuals 1 through 4 are half-sibs, as are 5 through 8, but unrelated to the first family. The relationship matrix \mathbf{A} becomes

$$\mathbf{A} = \begin{pmatrix} 1 & 0.25 & 0.25 & 0.25 & 0 & 0 & 0 & 0 \\ 0.25 & 1 & 0.25 & 0.25 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 1 & 0.25 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0.25 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0.25 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0.25 & 1 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0.25 & 0.25 & 1 & 0.25 \\ 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 1 \end{pmatrix}$$

These eight individuals are placed into two groups of size four. Group one contains individuals 1, 2, 5, 6, while group 2 contains 3, 4, 7, 8. For simplicity, the only assumed fixed-effect is the mean μ . The resulting matrices for the mixed-model equations become

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \\ z_8 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_d = \begin{pmatrix} A_{d,1} \\ A_{d,2} \\ A_{d,3} \\ A_{d,4} \\ A_{d,5} \\ A_{d,6} \\ A_{d,7} \\ A_{d,8} \end{pmatrix}, \quad \mathbf{Z}_d = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{I}_8$$

where $\beta = (\mu)$. The i th row in the incident matrix for social effects \mathbf{Z}_s has a one for each group member in i 's group, and a zero otherwise. For example, individual one is influenced by the social breeding values of individuals 2, 5, and 6, which are assigned values of one in the first row of \mathbf{Z}_s , while all other elements in row one are zero. Individual 3 is influenced by the social breeding values of its group – namely 4, 7 and 8. Filling in the rest of the matrix gives

$$\mathbf{Z}_s = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{a}_s = \begin{pmatrix} A_{s,1} \\ A_{s,2} \\ A_{s,3} \\ A_{s,4} \\ A_{s,5} \\ A_{s,6} \\ A_{s,7} \\ A_{s,8} \end{pmatrix}$$

The group effects are made more apparent (but \mathbf{A} more confusing) by reordering the individuals as $\mathbf{z}^T = (z_1, z_2, z_5, z_6, z_3, z_4, z_7, z_8)$, so that group members are clustered together. This gives

$$\mathbf{Z}_s = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}, \quad \mathbf{A} = 0.25 \cdot \begin{pmatrix} 4 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 4 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 4 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 4 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 4 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 4 \end{pmatrix}$$

Groups of different sizes are easily incorporated through the use of \mathbf{Z}_s .

Example 18.11. Assigning an element in \mathbf{Z}_s a value of one for each individual within a group weights all interactions equally. This need not be the case. For example, Muir (2006) notes that with plants (or other sessile organisms), the distance between two individuals likely influences their effects. In particular, he suggested that if trees are a distance d apart, a reasonable model for their associative interactions would be A_s/d^2 , where a tree has some intrinsic social breeding value A_s , whose effect is diminished by distance. In a case of three trees, where d_{ij} is the distance between trees i and j , the resulting incidence matrix for social breeding values would be

$$\mathbf{Z}_s = \begin{pmatrix} 0 & 1/d_{1,2}^2 & 1/d_{1,3}^2 \\ 1/d_{2,1}^2 & 0 & 1/d_{2,3}^2 \\ 1/d_{3,1}^2 & 1/d_{3,2}^2 & 0 \end{pmatrix}$$

Cantet and Cappa (2008) suggested similar “intensity of competition” weights for individuals within groups of animals, such as the total contact time between two individuals. Obviously, other similar measures could be used, and easily incorporated into \mathbf{Z}_s .

Since we allow for the possibility that the direct and social breeding values are correlated, the standard mixed-model equations for several vectors of random effects (Equation 15.21; LW Equations 26.19b, 26.30) must to slightly modified to allow this. They become

$$\begin{pmatrix} \mathbf{X}^T \mathbf{X} & \mathbf{X}^T \mathbf{Z}_d & \mathbf{X}^T \mathbf{Z}_s \\ \mathbf{Z}_d^T \mathbf{X}^T & \mathbf{Z}_d^T \mathbf{Z}_d + \lambda_1 \mathbf{A}^{-1} & \mathbf{Z}_d^T \mathbf{Z}_s + \lambda_2 \mathbf{A}^{-1} \\ \mathbf{Z}_s^T \mathbf{X}^T & \mathbf{Z}_s^T \mathbf{Z}_d + \lambda_2 \mathbf{A}^{-1} & \mathbf{Z}_s^T \mathbf{Z}_s + \lambda_3 \mathbf{A}^{-1} \end{pmatrix} \begin{pmatrix} \beta \\ \mathbf{a}_d \\ \mathbf{a}_s \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T \mathbf{X} \\ \mathbf{X}^T \mathbf{Z}_d \\ \mathbf{X}^T \mathbf{Z}_s \end{pmatrix} \quad (18.40a)$$

where the weights λ_i are related to elements in the inverse of \mathbf{G} , viz.,

$$\begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_2 & \lambda_3 \end{pmatrix} = \sigma_e^2 \mathbf{G}^{-1} = \sigma_e^2 \begin{pmatrix} \sigma^2(A_d) & \sigma(A_d, A_s) \\ \sigma(A_d, A_s) & \sigma^2(A_s) \end{pmatrix}^{-1} \quad (18.40b)$$

as obtained by Muir (2005) and Van Vleck and Cassady (2005).

In order to solve these equations, estimates for the variance components (\mathbf{G} and σ_e^2) are required, and within the mixed-model framework these are estimated by REML (LW Chapter 27). Van Vleck and Cassady (2005) used simulated data to show that, under the appropriate design, REML does indeed provide separable estimates of the genetic variance components. However, two early applications to real data sets, weight gain in pigs within pens by Arango et al. (2005) and weight gain in Hereford cattle in feedlots by Van Vleck et al. (2007) found that the likelihood surface for $\sigma^2(A_s)$ was very flat and that the model could not be fit. We will examine such issues of identifiability shortly. While mixed-model methodology is very robust (for example, easily handling missing data and variable group numbers), it can easily fail when the model is not correctly specified, a point we will also address shortly.

Example 18.12. Using mixed-model methods, Bergsma et al. (2008) examined four traits in domestic pigs (growth rate, feed intake, back fat thickness and muscle depth) in a sample of over 14,000 pigs with a known pedigree (from roughly 400 sires and 600 dams). Pigs were placed in pens of 6-12 animals and several different mixed models were fitted. First, a mixed-model allowing for only direct effects (plus a separate vector for common litter effects (\mathbf{c}), which is fairly standard) was run. Here is model is $\mathbf{z} = \mathbf{X}\beta + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_c\mathbf{c} + \mathbf{e}$, where $\mathbf{a} \sim N(0, \sigma_a^2 \mathbf{A})$, $\mathbf{c} \sim N(0, \sigma_c^2 \mathbf{I})$, and $\mathbf{e} \sim N(0, \sigma_e^2 \mathbf{I})$. The resulting estimates of additive variation and heritability for these traits was found to be

	Growth	Back fat	Muscle	Intake
$\sigma^2(A)$	2,583	2.83	7.94	41,275
h^2	0.37	0.36	0.25	0.41

Next, a model was fit that also included a random pen (group) effect g common to all members within the same group (but differing across groups). The model now becomes $\mathbf{z} = \mathbf{X}\beta + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_g\mathbf{g} + \mathbf{Z}_c\mathbf{c} + \mathbf{e}$, where $\mathbf{g} \sim N(0, \sigma_g^2 \mathbf{I})$. Use of this model did not change the heritabilities for back fat and muscle depth, but did decrease the estimates for growth and feed intake. The resulting estimates become

	Growth	Back fat	Muscle	Intake
$\sigma^2(A)$	1,780	2.79	7.69	17,678
h^2	0.25	0.36	0.24	0.18

Failure to include a group effect (here, assumed to be entirely non-heritable, i.e. all environmental), resulted in some traits having their heritabilities over-estimated. Finally, a model was fit allowing for heritable social effects, $\mathbf{z} = \mathbf{X}\beta + \mathbf{Z}_d\mathbf{a}_d + \mathbf{Z}_s\mathbf{a}_s + \mathbf{Z}_g\mathbf{g} + \mathbf{Z}_c\mathbf{c} + \mathbf{e}$, which gave estimates of

	Growth	Back fat	Muscle	Intake
$\sigma^2(A_d)$	1,522	2.75	6.68	16,950
h_d^2	0.21	0.35	0.21	0.17
$\sigma^2(A_s)$	51	0.01	0.03	596
$\sigma^2(A_T)$	5,208	3.19	10.35	68,687
T^2	0.71	0.41	0.32	0.70

While body-composition traits (back fat, muscle thickness) were largely unaffected by social effects, growth-related traits (growth, intake) were. Failure to incorporate group effects (either shared environmental g or genetic A_s) resulted in an overestimation of the heritability for growth traits. The exploitable genetic variance (the total breeding value) for the two growth traits was about three times higher than suggested by the individual (direct) breeding values, with T^2 about triple the value of h_d^2 . Hence, significant potential for improvement would remain untapped if individual selection with groups of unrelated individuals was used for growth traits. Conversely, incorporation of some group-selection would have little additional effect on the response of the two composition traits as h_d^2 and T^2 are similar for these traits.

The results in the above example are fairly typical of the published results from the animal breeding literature. Often the estimates of $\sigma(A_d, A_s)$ and $\sigma^2(A_s)$ are quite small relative to $\sigma^2(A_d)$, but since terms involving social effects are scaled by roughly n or n^2 (for the covariance and variance, respectively), their contributions can be considerable. For a series of eight (mostly growth) traits in cattle, pigs, and chicken, $(n-1)\sigma(A_d, A_s)$ was between 5 and 40% of $\sigma^2(A_d)$, with an average value of 24% (Van Vleck et al. 2007; Chen et al. 2008, 2009; Hsu et al. 2010).

As covered in Chapter 15, one could use a Bayesian analysis of a mixed model instead of BLUP estimates of the random effects and REML estimates of the variance. Recall that a BLUP/REML analysis returns point estimates and associated confidence intervals for variables of interest, while a Bayesian analysis returns the whole posterior distribution of potential values given the data (Chapter 15; Appendices 2, 3). Arora and Lahiri (1997) show for mixed models that “empirical BLUP”, namely using REML estimates of variance components to solve the mixed-model equations, generally gives the same average value as a Bayesian analysis, but that the later returns a smaller mean squared error and hence offers more precision. Cappa and Cantet (2008) develop a Gibbs sampler (Appendix 3) for the mixed-model with associative effects.

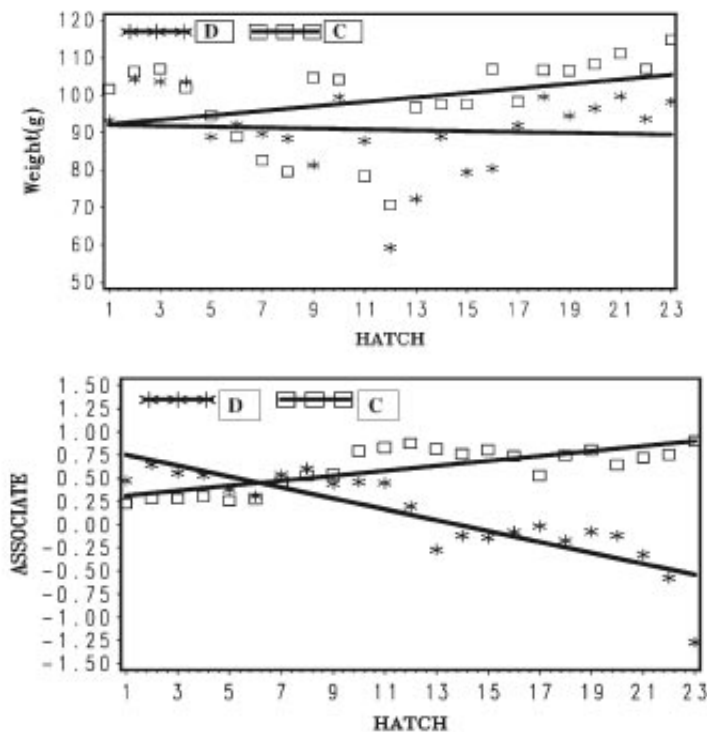


Figure 18.4. Selection response for two differentially-selected lines of Japanese quail. Both lines were selected for six-week weight using BLUP. Line D-BLUP selected individuals with the largest estimated direct breeding values, while line C-BLUP selected individuals with the largest estimated total breeding value. **A:** (Top). Mean response in six-week weight over 23 cycles of selection. The C-BLUP line showed a significant improvement, while the D-BLUP line showed a slight (but not significant) negative trend. **B:** (Bottom). The trend in mean social values showed an increase in the C-BLUP lines, and a decrease in D-BLUP lines. Hence, competition increased in lines strictly selected for direct breeding value, while it decreased in lines selected on an index of direct and associative effects.

Muir's Experiment: BLUP Selection for Quail Weight

In Muir's (2005) classic paper, he not only laid out the framework for incorporating social effects into a mixed-model framework, he also directly tested this approach by examining the response to selection based entirely on the estimated breeding values (EBVs) obtained

from the model. Muir selected on six-week weight in Japanese quail (*Coturnix coturnix japonica*), which are aggressive and cannibalistic. Groups were formed by placing 16 birds per cage, with each group consisting of several half-sib families. Banding of the birds allowed the pedigree of individuals to be followed through the 23 hatches of the experiment. As Example 18.4 shows, due to a negative covariance between associate and direct effects, individual selection is expected to produce a reversed response when using a group of unrelated individuals.

Rather than select using individual phenotype or group means, Muir used BLUP selection (Chapter 35), wherein a mixed model is used to estimate the breeding values, and those individuals with the largest EBVs are chosen. Starting with the same base population, two lines were selected using different BLUP criteria. For both lines, the mixed model allowing for both direct and social effects was fitted, using REML estimates of the variances to obtain BLUPs for the desired breeding values. In the D-BLUP line, individuals with the largest EBVs of A_d (direct effects) were selected. The C-BLUP line selected those individuals with the largest EBVs of A_T , namely $EBV(A_d) + (16 - 1)EBV(A_s)$. Figure 18.4A shows the results through 23 hatches (cycles of selection). Under BLUP-D selection, the mean six-week weight decreased (slightly, but not significant), while it significantly increased under C-BLUP. Both D-BLUP and C-BLUP increased the mean of direct effects, although the response under D-BLUP was about twice as great. As shown in Figure 18.4B, the reason for the decrease in mean weight in the D-BLUP line was that the associative effects increased under C-BLUP, but (as expected given the negative correlation between A_d and A_s) decreased under D-BLUP. Two other improvements were observed in the C-BLUP line. Morality increased significantly in the D-BLUP line, while decreasing slightly (but not significantly) in the C-BLUP lines. Feed conversion was also better in the C-BLUP line, requiring 6.65 grams of feed per gram of gain, versus a value of 7.26 in the D-BLUP line. Clearly, selection based on the mixed-model estimates of total breeding value resulted in significantly better results than lines selected by a more conventional (i.e., D-BLUP) approach.

Details: Environment Group Effects and the Covariance Structure of \mathbf{e}

The simplifying assumption that the residuals e_i are homoscedastic and uncorrelated (so that $\sigma(\mathbf{e}) = \sigma_e^2 \mathbf{I}$) is generally incorrect. As Equation 18.23a shows, individuals within the same group are correlated because they share the E_s values from the other group members. Not correctly accounting for these shared environmental values results in an overestimation of the variance of the social breeding values (Van Vleck and Cassady 2005, Bijma et al. 2007b, Bergsma et al. 2008, Chen et al. 2009). Equation 18.23a gives the correct covariance matrix for the residuals as

$$\sigma(\mathbf{e}) = \sigma_e^2 \mathbf{R}, \quad \text{where} \quad R_{ij} = \begin{cases} 0 & i \text{ and } j \text{ in different groups} \\ \rho & i \text{ and } j \text{ in the same group} \\ 1 & i = k \end{cases} \quad (18.41)$$

where σ_e^2 and ρ are given by Equation 18.23b.

Example 18.13. For the design used in Example 18.10 with group members clustered, so that $\mathbf{z}^T = (z_1, z_2, z_5, z_6, z_3, z_4, z_7, z_8)$, the corresponding covariance matrix for the residuals

becomes

$$\sigma(\mathbf{e}) = \sigma_e^2 \mathbf{R} = \sigma_e^2 \cdot \begin{pmatrix} 1 & \rho & \rho & \rho & 0 & 0 & 0 & 0 \\ \rho & 1 & \rho & \rho & 0 & 0 & 0 & 0 \\ \rho & \rho & 1 & \rho & 0 & 0 & 0 & 0 \\ \rho & \rho & \rho & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \rho & \rho & \rho \\ 0 & 0 & 0 & 0 & \rho & 1 & \rho & \rho \\ 0 & 0 & 0 & 0 & \rho & \rho & 1 & \rho \\ 0 & 0 & 0 & 0 & \rho & \rho & \rho & 1 \end{pmatrix}$$

With the same number of individuals in all groups, the only two estimable parameters in the environmental covariance matrix are ρ and σ_e^2 . With groups of variable size (either by design or simply through loss of data), the residual variances and covariances change with n (Equation 18.23b). In this cases, the residual covariance matrix would be specified in terms of the three environmental (co)variance terms, $\sigma^2(E_d)$, $\sigma^2(E_s)$, and $\sigma(E_d, E_s)$.

Provided $\rho > 0$, an equivalent approach is simply to fit a random group effect (Ellen et al. 2008, Bergsma et al. 2008). Example 18.14 works through an example. This approach is computationally less demanding than jointly estimating σ_e^2 and ρ in an \mathbf{R} matrix. However, if the covariance $\sigma(E_d, E_s)$ between environmental direct and social effects is sufficiently negative, ρ can be negative (Equation 18.23b) and the simple random group effects fails, as the group variance σ_g^2 must be positive. As Equation 18.23b suggests, as group size n increases, the contribution from $\sigma^2(E_s)$ eventually dominates ρ , making it positive. Thus, for a design with large group size, fitting a random group effect will usually suffice.

Example 18.14. Suppose that we instead of fully specifying \mathbf{R} , we simply fit a random group effect. Here, all individuals in group i show the common random effect g_i , where we assume $g_i \sim N(0, \sigma_g^2)$. The resulting mixed-model becomes

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_d\mathbf{a}_d + \mathbf{Z}_s\mathbf{a}_s + \mathbf{Z}_g\mathbf{g} + \mathbf{e}$$

where, for k groups, $\mathbf{g}^T = (g_1, g_2, \dots, g_k)$ is the vector of random group effects. The incident matrix \mathbf{Z}_g has k columns, the i th of which (corresponding to membership in group i) has a one for each individual in group i and a zero elsewhere. For example, for the design in Example 18.13,

$$\mathbf{g} = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}, \quad \mathbf{Z}_g = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$$

Assuming a simple covariance for the residuals, so that $\sigma(\mathbf{e}) = \sigma_e^2 \mathbf{I}$, the resulting contribution

to the covariance matrix of \mathbf{z} from the group and residual terms becomes $\sigma_g^2 \mathbf{Z}_g \mathbf{Z}_g^T + \sigma_e^2 \mathbf{I}$, or

$$\sigma_g^2 \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} + \sigma_e^2 \mathbf{I} = \sigma_e^2 \begin{pmatrix} 1 & \alpha^2 & \alpha^2 & \alpha^2 & 0 & 0 & 0 & 0 \\ \alpha^2 & 1 & \alpha^2 & \alpha^2 & 0 & 0 & 0 & 0 \\ \alpha^2 & \alpha^2 & 1 & \alpha^2 & 0 & 0 & 0 & 0 \\ \alpha^2 & \alpha^2 & \alpha^2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \alpha^2 & \alpha^2 & \alpha^2 \\ 0 & 0 & 0 & 0 & \alpha^2 & 1 & \alpha^2 & \alpha^2 \\ 0 & 0 & 0 & 0 & \alpha^2 & \alpha^2 & 1 & \alpha^2 \\ 0 & 0 & 0 & 0 & \alpha^2 & \alpha^2 & \alpha^2 & 1 \end{pmatrix}$$

where

$$\sigma_e^2 = \sigma_e^2 + \sigma_g^2, \quad \alpha^2 = \frac{\sigma_g^2}{\sigma_e^2 + \sigma_g^2}$$

We use α^2 to remind the reader that (under this model), this is the ratio of two variances, and hence is always non-negative. Comparison with Example 18.13. shows that adding a random effect for group corresponds to the more fully-specified covariance residual model (Equation 18.41), with $\rho = \alpha^2$. When $\rho > 0$, both models are identical. However, if the within-group environmental correlations are negative ($\rho < 0$), then the simple group random effect model fails.

Finally, note that we can write the covariance matrix of group plus residual effects under the model as

$$\sigma_e^2 \left(\mathbf{I} + \alpha^2 [\mathbf{Z}_g \mathbf{Z}_g^T - \mathbf{I}] \right)$$

show that \mathbf{R} in Equation 18.41 (provided $\rho > 0$) is given by

$$\mathbf{R} = \mathbf{I} + \alpha^2 [\mathbf{Z}_g \mathbf{Z}_g^T - \mathbf{I}]$$

Finally, we should point out that a standard approach when families are in the analysis is to include a common family effect c due to shared maternal effects and/or dominance (if full sibs are present). This is simply done by adding an additional random effect c for the family effects. For our model in this example, this becomes

$$\mathbf{z} = \mathbf{X}\beta + \mathbf{Z}_d \mathbf{a}_d + \mathbf{Z}_s \mathbf{a}_s + \mathbf{Z}_g \mathbf{g} + \mathbf{Z}_c \mathbf{c} + \mathbf{e}$$

Suppose that individuals 1,2,5,6 are from one family, while the others are from a second. Here

$$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \quad \mathbf{Z}_c = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$$

Note that if $\mathbf{Z}_g = \mathbf{Z}_c$, the group and family effects are fully confounded and cannot be separated (also see below). This additional of a common family effect is easily incorporated into any of our above models.

Details: Ignoring Additive Social Values Introduces Bias

Before models directly accounting for social effects, it was not unusual to add a fixed or random group effect to the standard animal model to account for common environments due to being raised in pens, cages, or groups. For example, assuming group effects are random, the corresponding animal model becomes

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{Z}_g\mathbf{g} + \mathbf{e} \quad (18.42)$$

In this model, \mathbf{a} would be the estimated vector of (direct) breeding values. As detailed above, \mathbf{g} can often account for any shared environmental social values (i.e., E_s). However, if additive associate effects are present, simply adding a group effect is insufficient, resulting in overestimation of σ_g^2 and often overestimates the (direct) additive variance (Example 18.12). Hence, an analysis that simply includes a group effect (but no \mathbf{a}_s) results in biased estimates of the direct breeding values. Van Vleck and Cassady (2005) show how the presence of additive associative effects inflates the estimate of group variance. Consider two members in the same group (with common group effect g),

$$\begin{aligned} z_1 &= A_{d_1} + A_{s_2} + \sum_{k=3}^n A_{s_k} + g + e_1 \\ z_2 &= A_{d_2} + A_{s_1} + \sum_{k=3}^n A_{s_k} + g + e_2 \end{aligned}$$

Using the standard ANOVA identity that the covariance within a group equals the variance between groups, for unrelated individuals σ_g^2 is estimated from the within-group covariance, which reduces to

$$\begin{aligned} \sigma(z_1, z_2) &= \sigma(A_{d_1}, A_{s_1}) + \sigma(A_{d_2}, A_{s_2}) + (n-2)\sigma^2(A_s) + \sigma_g^2 + \sigma(e_1, e_2) \\ &= 2\sigma(A_d, A_s) + (n-1)\sigma^2(A_s) + \sigma_g^2 + \sigma(e_1, e_2) \end{aligned} \quad (18.43)$$

Assuming the residuals are uncorrelated, the bias in estimation of the common group variance σ_g^2 is $2\sigma(A_d, A_s) + (n-1)\sigma^2(A_s)$, which can be considerable. Hence, when additive associative effects are present, the simple model given by Equation 18.42 is inappropriate. This model, however, can be useful in a preliminary analysis. Van Vleck and Cassady suggest that obtaining a large estimated group variance when using Equation 18.42 indicates that a more detailed model including additive associative effects should be fit to the data. Hence, one approach is to do a quick fit to Equation 18.42. If the group variance is sufficiently small, it is unlikely that additive associative effects are present. This approach is not always fool-proof, as inspection of Equation 18.43 shows that a sufficiently negative covariance between direct and social breeding values may result in a small estimated group variance.

Details: Identifiability of Variance Components

Due to potential confounding of effects, any particular design might not allow for all variables of interest to be uniquely estimated. For the vector $\boldsymbol{\beta}$ of fixed effects, this is the concept of **estimability** (LW Chapter 26). For $\mathbf{z} \sim (\mathbf{X}\boldsymbol{\beta}, \mathbf{V})$, the vector of fixed effects is estimable (all have unique values) if $(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}$ exists. Otherwise, some of the fixed effects are confounded and cannot be separated by the design (\mathbf{X}) being used. With (co)variance components (often called **dispersal parameters**), a similar concept, **identifiability**, also exists. If variance components are not identifiable in the design, then BLUPs for their associated vectors of random effects do not exist.

Lack of identifiability has been a problem in some studies, with lack of convergence of REML estimates, convergence to multiple peaks in the likelihood surface (depending on starting conditions), and very flat likelihood surfaces all seen (Arango et al. 2005, Van Vleck et al. 2007, Chen et al. 2008). Cantet and Cappa (2008) formally showed that using a *fixed* group effect results in lack of identifiability when the design matrix \mathbf{Z}_g contains equal weights for all group members. Thus treating group effects as fixed is not recommended, while treating them as random can often account for environmental correlations (as discussed above). Another common source of lack of identifiability is the composition of the group. If all group members are from a single half-sib or full-sib family, the covariance of group members equals the covariance among family members within a group, confounding variance components and leading to a lack of identifiability. Bijma et al. (2007b) noted that this problem plagued one of the early attempts to estimate social variance components (Wolf 2003). The important caveat is that lack of identifiability can easily arise in attempts to estimate social effects even when using seemingly innocent designs (such as a fixed group effect or having each group be a single family).

Conditions for identifiability of REML estimates of (co)variance components are given by Rothenberg (1971), Jiang (1996), and Cantet and Cappa (2008). Before presenting these, we first review a few details about REML. Recall (LW Chapter 27) that REML estimates are those that maximize that part of the likelihood function that is independent of the fixed effects (this is often stated as being the **translation invariant** part). Let \mathbf{V} be the covariance matrix of \mathbf{z} , which is a function of its variance components. As detailed in LW Chapter 27, Harville (1977) shows that (if it exists) the transformation provided by the matrix

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1} \quad (18.44a)$$

plays a critical role in REML estimates. That this matrix can remove fixed effects can be seen by noting that

$$\mathbf{P}\mathbf{z} = \mathbf{V}^{-1}(\mathbf{z} - \mathbf{X}\hat{\boldsymbol{\beta}}) \quad (18.44b)$$

yields a vector that is the data vector adjusted by the (estimated) fixed effects. Now consider covariance structures of the form

$$\mathbf{V} = \sum_{i=1}^n \mathbf{V}_i \theta_i \quad (18.45a)$$

where \mathbf{V}_i is a matrix of known constants and the θ_i are unknown variances and covariances to be estimated. The equations to maximize the likelihood over the restricted space (the REML estimates) are given by LW Equations 27.18 and 27.19, and are solved iteratively. These equations involve the **trace** (sum of the diagonal elements) of matrix products involving \mathbf{P} and the \mathbf{V}_i . Recall (LW Appendix 4) that for a vector $\boldsymbol{\theta}$ of n unknowns, the Fisher information matrix \mathbf{F} (the matrix of second partial derivatives of the likelihood with respect to the parameters) can be used to provide large-sample standard errors. The resulting $n \times n$ information matrix for REML estimates of the unknown θ_i in Equation 18.45a is

$$F_{ij} = \text{trace}(\mathbf{P}\mathbf{V}_i\mathbf{P}\mathbf{V}_j) \quad (18.45b)$$

Much in the same fashion that the existence of $(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}$ informs us that all fixed effects are estimable in a given design, all variance components θ_i are identifiable if all of the eigenvalues of \mathbf{F} are positive, that is, that \mathbf{F} is positive-definite (Rothenberg 1971, Jiang 1996). For the simplest associative effects mixed model (Equation 18.38), Equation 18.45a becomes

$$\mathbf{V} = \mathbf{V}_1 \sigma^2(A_d) + \mathbf{V}_2 \sigma(A_d, A_s) + \mathbf{V}_3 \sigma^2(A_s) + \mathbf{V}_4 \sigma_e^2 \quad (18.46a)$$

where

$$\mathbf{V}_1 = \mathbf{Z}_d \mathbf{A} \mathbf{Z}_d^T, \quad \mathbf{V}_2 = (\mathbf{Z}_d \mathbf{A} \mathbf{Z}_s^T + \mathbf{Z}_s \mathbf{A} \mathbf{Z}_d^T), \quad \mathbf{V}_3 = \mathbf{Z}_s \mathbf{A} \mathbf{Z}_s^T, \quad \mathbf{V}_4 = \mathbf{I} \quad (18.46b)$$

Substituting Equations 18.44a and 18.46b into Equation 18.45b fills out the \mathbf{F} matrix (which is only 4×4 in this case given the four unknown variance components). For any particular design, the eigenvalues of this matrix can be computed to determine if the variance components are all identifiable.

Final Thoughts: Appropriate Designs for Estimating Direct and Associative Effects

In contrast to selection response where there is benefit from having all group members from the same family (and hence an increased r value), in a design to estimate direct and associative values and variance components, groups should be composed of at least two different families. Groups can also consist of unrelated individuals, but Bijma (2010c) shows that, in general, using groups with two different families offers more power than using unrelated individuals.

Using the appropriate model is also critical. Initially, one might think that associative effects could be accounted for by simply adding a random effect for group to an otherwise standard animal model. If genetic associative effects are present, this typically over-estimates the direct effects as well as inflating the group variance (which is a measure of the environmental social effects). One use of adding a simple group effect is as a quick check for potential social effects. If the resulting estimate of the group variance is modest to large, a full model should be fitted. However, if the group variance estimate is small, then the contribution for associative effects is likely (but not guaranteed!) to be small as well.

Conversely, ignoring any environmental associative effects also introduces bias. For example, a model fitting just \mathbf{a}_d and \mathbf{a}_s with a simple error structure $e \sim (0, \sigma_e^2 \mathbf{I})$ also introduces bias. As mentioned above, the correct residual covariance structure can be accounted for by incorporating a random group effect into the model (which assumes a positive correlation between social environmental effects within a group), or by using a model with $e \sim (0, \sigma_e^2 \mathbf{R})$ where the elements of \mathbf{R} are given by 18.41, which allows for the within-group environmental correlations ρ to be negative.

ASSOCIATIVE-EFFECTS, INCLUSIVE FITNESS, AND FISHER'S THEOREM

We conclude the chapter by briefly examining some of the important implications for evolution when heritable associative effects ($\sigma^2(A_s) > 0$) are present. First and foremost, their presence has significant implications on the evolution of mean population fitness (Bijma 2010a), which forms the subject of this section.

Change in Mean Fitness When Associative Effects are Present

The most important trait in evolution is fitness, W . Clearly, the fitness of any particular genotype is partly a function of the environment in which it finds itself. While we normally treat this environment as static, when conspecifics influence fitness (as is generally expected to be the case!), part of this environment may also be evolving in response to selection. In these settings, models with associative effects are appropriate. Here the individual fitness of a focal individual results from a direct fitness effect from its own genotype plus the associative effects on its fitness from the other genotypes with which it interacts. Competition, a detrimental fitness effect from other individuals, is one such associative effect, where the contribution from conspecifics is to lower fitness. With cooperation or mutualism, associative effects increase the fitness of the focal individual.

Examining the expected change in mean fitness is straightforward. Using previous results, we simply take the trait being followed as individual fitness ($z = W$). From Equation

18.1c, the fitness of individual i becomes

$$W_i = \mu + A_{d_i} + \sum_{j \neq i} A_{s_j} + e_i \quad (18.47a)$$

A_d is the direct breeding value of fitness, while A_s is the social breeding value (how a focal individual influences the fitness of others in its group). As above, A_{s_i} does not contribute to W_i , while A_{s_j} for $j \neq i$ does. Likewise, as before the total breeding value for fitness of an individual is simply

$$AT_i = A_{d_i} + (n-1)A_{s_i} \quad (18.47b)$$

with variance

$$\sigma^2(AT) = \sigma^2(A_d) + 2(n-1)\sigma(A_d, A_s) + (n-1)^2\sigma^2(A_s) \quad (18.47c)$$

The first term is the classical additive genetic variance in fitness in the absence of associative effects. When interactions are present, there is the potential for substantially more heritable variation in fitness. Indeed, the total genetic variance in fitness has the potential to exceed the actual variance in individual fitness ($\sigma^2(AT) > \sigma^2_W$), as much of the variation is hidden in interactions with others, which do not appear in one's individual fitness.

When the trait is fitness itself, the response equation for individual selection (Equation 18.10) simplifies somewhat. Recall the Robertson-Price identity (Equation 10.7), $S = \sigma(z, w)$, where $w = W/\bar{W}$ is relative fitness. When $z = W$, the selection differential becomes

$$S_W = \sigma(W, w) = \frac{\sigma(W, W)}{\bar{W}} = \frac{\sigma^2(W)}{\bar{W}} \quad (18.48a)$$

Substituting into Equation 18.10 gives the response (the change in mean population fitness) as

$$R_W = \frac{\sigma(W, AT)}{\sigma^2(W)} S_W = \frac{\sigma(W, AT)}{\sigma^2(W)} \frac{\sigma^2(W)}{\bar{W}} = \frac{1}{\bar{W}} \sigma(W, AT) \quad (18.48b)$$

This is simply Price's Equation (Equation 12.1), where we have restricted our discussion to cases where the transmission is such that mean breeding values of offspring equals the mean breeding values of their parents (and hence the correction term in Equation 12.1 for changes induced solely by transmission disappears). Such is expected to be the case for the infinitesimal model under random mating.

Applying Equation 18.12c gives the response in terms of the variance components as

$$R_W = \frac{1}{\bar{W}} [\sigma^2(A_d) + (n-1)(1+r)\sigma(A_d, A_s) + r(n-1)^2\sigma^2(A_s)] \quad (18.48c)$$

Just as we have seen for other traits, when $r = 0$, the possibility of a reversed response occurs if the breeding value for direct and social effects on fitness are sufficiently negatively correlated. Hence, under rather realistic conditions, individual selection can result in a *decrease* (and a potentially rather significant one at that) in mean fitness.

Example 18.15. As Equation 18.48c shows, a reversed response in mean population fitness can occur, in the extreme driving a population to extinction. A potential example of this is Dawson (1969), who was studying competition between two species of flour beetles (*Tribolium castaneum* and *T. confusum*). He found that *castaneum* won (driving the other species to extinction) in nine of ten replicates. In the remaining replicate, *castaneum* again appeared

to be winning (with a frequency of over 90% by generation four from a starting frequency of total individuals of 50%) when an eye color mutation allelic to *chestnut* spontaneously appeared. From that generation onward, the frequency of this allele increased while the population itself was driven to extinction. One explanation for such a **Trojan gene** (Muir and Howard 1999) is a strong positive direct fitness effect (increasing the frequency of the allele), but with a strong negative associative value on conspecifics (decreasing mean population fitness).

Ironically, even though a negative response can occur in the presence of associative effects, there is actually more total variance potentially available when they are present, as $\sigma^2(A_T) \geq \sigma^2(A_d)$. However, only a fraction of this may be accessible to *individual* selection, and this fraction (being a covariance rather than a variance) can be negative. The key for exploiting the available variance is either selection among groups and/or the presence of relatives in one's group of interacting individuals.

To see this, note from Equation 18.12e that we can express Equation 18.48c as

$$R_W = \frac{1}{\bar{W}} \left(r\sigma^2(A_T) + (1-r) [\sigma^2(A_d) + (n-1)\sigma(A_d, A_s)] \right) \quad (18.48d)$$

The term in square brackets represents the response in a group of non-relatives. When interactions occur among kin ($r > 0$), then for sufficiently close relatives, the response becomes positive (mean fitness increases) even if it is negative when $r = 0$. At the extreme, when $r = 1$ (all interactions are among clones), the response in mean fitness is simply $\sigma^2(A_T)/\bar{W}$ and all of the heritable variance in fitness is utilized. Conversely, when interactions occur among unrelated individuals, only a fraction of this genetic potential is exploited. This observation lead Bijma (2010a) to suggest that when heritable fitness interactions are present, the key to evolutionary success is interacting with relatives. The reason for this is clear from our previous discussions. With interactions among unrelated individuals, one's phenotype (here fitness) provides very little information about their social breeding value. With interacting kin, the breeding values of the kin's social effects influences your fitness, and these are positively correlated (via kinship) with your own breeding value for social effects.

Finally, we can decompose the total response in fitness into response from changes in the mean of the direct effects and response from changes in the mean of the social effects. Equation 18.15a gives

$$R_W = R_{W,d} + (n-1)R_{W,s} \quad (18.49a)$$

Recalling Equation 18.48a, Equations 18.15b,c give these response components as

$$R_{W,d} = \frac{\sigma^2(A_d) + r(n-1)\sigma(A_d, A_s)}{\bar{W}} \quad (18.49b)$$

and

$$R_{W,s} = \frac{\sigma(A_d, A_s) + r(n-1)\sigma^2(A_s)}{\bar{W}} \quad (18.49c)$$

Example 18.16. Haldane (1932) coined the term **altruistic trait** to denote a behavior (or trait) that harms an individual, but benefits others. The classic example are alarm calls — others in a group are warned (increasingly their fitness), but at some expense to the individual making the call (a direct effect decreasing fitness). Note that the increase in an altruistic trait is an

example of a reversed response, as the trait lowers the fitness of the individual that bears it. What are the conditions for such traits to spread? In terms of our fitness model with associative effects (Equation 18.47a), we can rephrase this as the conditions for the mean value of A_s to increase, which are given by Equation 18.49c. From the definition of altruism, $\sigma(A_d, A_s) < 0$, as performing an altruistic act decreases your direct fitness while increasing the fitness of those in your group. Equation 18.49c shows that a necessary (but not sufficient!) condition for altruism to evolve under individual selection is $r > 0$, i.e., individuals interact in groups of relatives.

As pointed out by Bijma and Wade (2008), we can view $\sigma(A_d, A_s)$ as the **cost** ($-c$) for an altruistic act towards others in your group. Conversely, the altruistic contribution to you from others in your group is $(n-1)\sigma^2(A_s) \geq 0$, which we denote as the **benefit** b . With these definitions, from Equation 18.49c the condition for altruism to evolve under individual selection is just

$$-c + rb > 0, \quad \text{or} \quad r > b/c$$

This is the classic **Hamilton's rule** (Hamilton 1963; 1964a,b), which we discuss shortly. For altruism to evolve under individual selection requires that individuals interact in groups of relatives. If groups consist of unrelated individuals, individual selection is not sufficient for altruism, and some component of group selection is required. Note that a more general version of Hamilton's rule is

$$\sigma(A_d, A_s) + r(n-1)\sigma^2(A_s) > 0,$$

which shows that the benefit scales with group size, provided that $\sigma^2(A_s)$ is independent of n . In contrast, under Bijma's dilution model (Equation 18.10a), the variance in social effects is a function of n , and (applying Equation 18.10c) Hamilton's rule becomes

$$\sigma(A_d, A_s) + r(n-1)^{1-2d}\sigma^2(A_{s,2}) > 0$$

with benefit increasing with group size only when the dilution fraction $d < 1/2$.

The same logic extends to the evolution of altruism under group selection. From Equation 18.27a, the response in the mean social value to group selection depends on the covariance of A_s and the group mean (here the total group fitness). From Equation 18.27c,

$$\begin{aligned} \sigma\left(A_{s_i}, \sum_{j=1}^n W_j\right) &= [\sigma(A_d, A_s) + (n-1)\sigma^2(A_s)] [1 + (n-1)r] \\ &= (-c + b) [1 + (n-1)r] \end{aligned}$$

As long as $b > c$ (the benefit exceeds the cost), altruism can evolve under group selection, even in groups of *unrelated* individuals, as the sign of the covariance is independent of r . When individuals within the group are related, the covariance is significantly larger and hence the response to selection greater.

Inclusive Fitness

As Equation 18.47a illustrates, when heritable interactions are present, the fitness of an individual depends on both their own genes as well as the genes in others. Hamilton (1964a,b) suggested that the focus should shift from individual fitness to what he called **inclusive fitness** — that component of fitness influenced only by the alleles carried by the focal individual. Hamilton argued that individuals strive to increase their inclusive, as opposed to individual, fitness (also see Michod and Abugov 1980, Grafen 2006). Formally, the inclusive fitness of an individual is context-specific, and is defined as individual fitness minus any

contribution to that fitness from the group environment plus the effect of that individual on the fitness of others, weighted by relatedness. While sounding rather abstract, when placed in an associative effect framework, this definition is quite clear. From Equation 18.47a, for individual i , A_{d_i} is the heritable component of individual fitness W_i remaining when the social contributions from others have been removed. The focal individual's social breeding value A_{s_i} does not influence their own fitness, but the social effects of other group members do, with the (heritable) contribution to individual i 's fitness from individual j being A_{s_j} . The correlation between the breeding value A_{s_i} carried by i and the contribution to i 's fitness from j is their relatedness r_{ij} , so that $r_{ij}A_{s_i}$ is the predicted value of A_{s_j} given A_{s_i} . Putting these together gives the heritable component (i.e., breeding value) of i 's inclusive fitness as

$$A_{incl,i} = A_{d_i} + A_{s_i} \sum_{j \neq i}^n r_{ij} = A_{d_i} + r(n-1)A_{s_i} \quad (18.51a)$$

where the last equality makes our standard assumption that all group members are equally related (which is easily relaxed). The resulting variance in the breeding value for inclusive fitness becomes

$$\sigma^2(A_{incl}) = \sigma^2(A_d) + 2r(n-1)\sigma(A_d, A_d) + r^2(n-1)^2\sigma^2(A_s) \quad (18.51b)$$

In the absence of heritable associative effects ($\sigma^2(A_s) = 0$) this simply reduces to the additive variance in direct fitness. Importantly, note that the heritable component of inclusive fitness is not the same as the total breeding value A_T for fitness, as

$$A_{T_i} = A_{incl,i} + (1-r)(n-1)A_{s_i} \quad (18.51c)$$

Just as Equation 18.49 decomposed the total response into components from direct and associative effects, we can similarly decompose the change in mean individual fitness into change in mean inclusive fitness plus the residual response. From Equation 18.51c,

$$R_W = R_{W,incl} + (1-r)(n-1)R_{W,s} \quad (18.52a)$$

so that total response in fitness is the change in inclusive fitness plus any response in the residual of the mean social value (after the effects of group relatives are absorbed into inclusive fitness). From Equation 18.48b, the response in the mean inclusive fitness is given by

$$R_{W,incl} = \frac{1}{\bar{W}} \sigma(W, A_{incl}) \quad (18.52b)$$

where

$$\begin{aligned} \sigma(W, A_{incl}) &= \sigma \left(\mu + A_{d_i} + \sum_{j \neq i} A_{s_j} + e_i, A_{d_i} + r(n-1)A_{s_i} \right) \\ &= \sigma^2(A_d) + 2r(n-1)\sigma(A_d, A_d) + r^2(n-1)^2\sigma^2(A_s) \end{aligned} \quad (18.52c)$$

The last line follows by evaluating the covariance in a similar fashion as done throughout this chapter. Note by comparison with Equation 18.51b, that this is simply $\sigma^2(A_{incl})$, yielding

$$R_{W,incl} = \frac{\sigma^2(A_{incl})}{\bar{W}} \quad (18.52d)$$

Hence (under our simple model), the response in mean inclusive fitness is proportional to the additive variance in inclusive fitness, so that mean inclusive fitness is non-decreasing. Why, then, can the mean of *individual* fitness decline despite the continual increase in mean inclusive fitness? The reason is an even faster decline in the mean (residual) social value. Recalling Equation 18.49c, Equation 18.51a becomes

$$R_W - R_{W, incl} = \frac{(1-r)(n-1)}{\bar{W}} \left(\sigma(A_d, A_s) + r(n-1)\sigma^2(A_s) \right) \quad (18.52e)$$

Hence, if the covariance between direct and associative effects is sufficiently negative, any increase in inclusive fitness is more than countered by the decline in the mean social environment. Note that increasingly the relatedness of group members decreases the residual response between mean individual and inclusive fitness, which in turn increases the chances that individual mean fitness increases.

Bijma's Theorem: Inclusive Fitness and Fisher's Fundamental Theorem

As we have seen, when heritable associative effects are present, individual selection can cause mean individual fitness to decrease. Further, while mean *individual* fitness can decrease, mean *inclusive fitness* is nondecreasing, as the response in inclusive fitness is proportional to its additive variance (Equation 18.52d). Comparing these results with those from Chapter 7, we have an apparent conflict. For the simple model of additive fitness effects and random mating, the classical interpretation of Fisher's Fundamental Theorem (FFT) holds, with the change in mean individual fitness being proportional to the additive variance in individual fitness, so that (under these conditions) mean individual fitness is non-decreasing. Further, the Price-Ewens exact version of the FFT states that the partial increase in mean individual fitness (the change that occurs when set in a constant environment) is nondecreasing. Bijma (2010a) reconciled these issues, showing that the Price-Ewens FFT corresponds to statements about inclusive fitness (which reduces to individual fitness in the absence of associative effects).

Recall from Chapter 7 that Price (1972) and Ewens (1989b, 1992) showed that Fisher appeared not to be concerned about the total change in individual fitness, but rather only one component of that change, namely that due to changes in the allele frequencies of genes under selection, when all other factors (such as change in the environment) are held constant. Bijma (2010a) makes the important distinction between what we have been calling the total breeding value A_T and the breeding value BV computed by only considering the effects of the alleles in the focal individual on its own phenotype. The later (BV) is the more traditional definition of breeding value and (LW Chapter 4) is simply the regression of fitness on the direct and social breeding values of individual i ,

$$W_i = BV_i + e = \beta_{i,d}A_{d_i} + \beta_{i,s}A_{s_i} + e \quad (18.53a)$$

To compute these regression slopes, first note that individual fitness of individual i is a function of its direct value plus the associative effects for every group member except i ,

$$W_i = \mu + A_{d_i} + \sum_{j \neq i} A_{s_j} + e_i$$

Recalling that $\sigma(A_{s_i}, A_{s_j}) = r_{ij}\sigma^2(A_s)$, i 's "traditional" breeding value for individual fitness reduces to

$$BV_i = A_{d_i} + A_{s_i} \sum_{j \neq i} r_{ij} \quad (18.53b)$$

as obtain be Bijma (2010a). For the case of all relatives being equally related, this reduces to

$$BV_i = A_{d_i} + r(n-1)A_{s_i} \quad (18.53c)$$

Comparing this to Equation 18.51a shows that the standard breeding value BV equals the breeding value for inclusive fitness. Thus, we have **Bijma's theorem**: The traditional breeding value for individual fitness is equivalent to the breeding value for inclusive fitness. From Equation 18.52d, it immediately follows that the response in inclusive fitness is proportional to the traditional additive variance in fitness (which in this case is the variance in inclusive fitness). Hence, inclusive fitness provides the bridge between the FFT and response when associative effects are present.

HAMILTON'S RULE

Example 18.16 briefly introduced Hamilton's classic result: the condition for an altruistic trait to spread is $br - c > 0$, where c is the fitness cost to the individual (or **actor**) performing the altruistic act, b the benefit to individuals with which it interacts, and r is the relationship between the actor and the recipient (Hamilton 1963, 1964a,b). Hamilton's result followed by considering an altruistic trait determined by a single locus with additive fitness effects in an outbred population. In this setting, $-c + rb$ is the inclusive fitness of the altruistic allele, whose frequency increases when its inclusive fitness is positive. He further showed (again under this simple model) that inclusive, rather than individual, fitness is maximized by selection. The roots of Hamilton's result, that for altruism to spread under individual selection requires interactions occur among sufficiently related individuals, dates back to Haldane (1955). The widely attributed quote by Haldane (which is not in his 1955 paper) is "Would I lay down my life to save my brother? No, but I would to save two brothers or eight cousins." Hamilton's contribution was to generalize Haldane's intuition and to introduce the important concept of inclusive fitness.

How General is Hamilton's Rule?

Hamilton's rule is a bit like the breeder's equation: it provides a simple expression that conveys the nature of interactions between the key quantities of interest, but it can fail (at least as an exact expression) under a variety of conditions. Given that Hamilton assumed a single additive locus under weak selection in an outbred population, just how general is his result? Assuming weak selection, a number of studies have shown that altruistic traits with a polygenic basis also generally satisfy Hamilton's rule (e.g., Yokoyama and Felsenstein 1978, Boyd and Richerson 1980, Aoki 1982, Engels 1983). Conversely, Cavalli-Sforza and Feldman (1978) found that it can fail for even a single locus. When fitnesses are no longer additive (i.e., the fitness of the heterozygote is no longer the average of the two homozygotes), then comparing r to c/b may not be sufficient to predict spread or loss, as the actual value of b can also matter. Starting with Hamilton (1970), attempts for a more general version of Hamilton's rule were built around the Price Equation (Chapter 12). Recall from Equation 12.1a that Price's (1970, 1972a) theorem states that the response in any quantity G can be expressed as

$$\Delta G = \sigma(G, w) + E[w \delta_G] \quad (18.54)$$

where w is relative fitness and δ_G is any transmission bias, namely the average deviation between the value of G in an ancestor and its mean value in their descendants. Typically, we can treat G as either the frequency of an allele (for a single-locus analysis) or as the breeding value of a trait (for a quantitative-genetic analysis). In the absence of drift and under normal Mendelian segregation (i.e., no meiotic drive), the allele frequencies in the offspring match the allele frequencies in their parents. Likewise, under the infinitesimal model (in an outbred population) the expected breeding value of the offspring equals the mean breeding value of its parents, and hence the second term in Price's Equation is usually ignored (see Chapter

12 for a much fuller discussion of this transmission term).

One way to obtain Hamilton's result is to assume what is typically (in the social evolution literature) called **neighbor-modulated fitness** (Hamilton 1964a,b), namely the phenotype(s) of your neighbor(s) influences your fitness (more recently these have been referred to as **direct fitnesses**, e.g., Taylor and Frank 1996, Taylor et al. 2007). Following Queller (1992a), let the fitness of some focal individual i interacting with its neighbor j be

$$W_i = W_0 + W_{d,i} + W_{s,j} \quad (18.55a)$$

W_0 is a component independent of social interactions, $W_{d,i}$ is the direct effect of i on its own fitness, and $W_{s,j}$ the (social) effect of the neighbor j on i 's fitness. Now, use a linear regression to quantify the effects of phenotypes z_i on fitness, with

$$W_i = a + \beta_{W_d|z} z_i + \beta_{W_s|z} z_j + e_i \quad (18.55b)$$

Ignoring the transmission bias term in Equation 18.54, Price's theorem gives the expected response in G as

$$\begin{aligned} \Delta G = \sigma(W_i, G_i) &= \sigma(\beta_{W_d|z} z_i + \beta_{W_s|z} z_j + e_i, G_i) \\ &= \beta_{W_d|z} \sigma(z_i, G_i) + \beta_{W_s|z} \sigma(z_j, G_i) + \sigma(e_i, G_i) \end{aligned} \quad (18.55c)$$

Here, we have used the fact that $\sigma(a, G_i) = 0$ for the constant a . Note that, to this point, Equation 18.55c is exact. If the residual e_i from the regression of fitness on phenotype (Equation 18.55b) is uncorrelated with the genotypic value G_i , we have exactly that

$$\Delta G = \beta_{W_d|z} \sigma(z_i, G_i) + \beta_{W_s|z} \sigma(z_j, G_i) \quad (18.56a)$$

The residual condition is satisfied when W is entirely determined by the phenotypes z_i and z_j plus any additional components *provided* they are uncorrelated with G_i . The trait will spread when $\Delta G > 0$. Dividing through by $\sigma(z_i, G_i)$, Equation 18.56a gives this condition as

$$\beta_{W_d|z} + \beta_{W_s|z} \frac{\sigma(z_j, G_i)}{\sigma(z_i, G_i)} > 0 \quad (18.56b)$$

where $\beta_{W_d|z}$ is the cost c of the trait, $\beta_{W_s|z}$ the benefit b , and the ratio is a more generalized measure of relatedness. Given that the phenotypic fitness regression residuals are uncorrelated with breeding values (and that we can safely ignore the transmission bias term), then Equation 18.56b is a general version of Hamilton's rule.

Under the infinitesimal model assumptions in a non-inbred population, the regression term (a covariance divided by a variance) quantifying relatedness reduces to

$$\frac{\sigma(z_j, G_i)}{\sigma(z_i, G_i)} = \frac{\sigma(G_j, G_i)}{\sigma(G_i, G_i)} = \frac{r \sigma(G_i, G_i)}{\sigma(G_i, G_i)} = r$$

and we recover the standard version of Hamilton's rule. However, when inbreeding is present and/or the infinitesimal model assumptions are not valid (i.e., significant allele frequency change can occur within a generation), Price's Equation shows that a more general definition of relatedness may be required. There is a detailed literature on the appropriate measure of relatedness to use, see Hamilton (1970), Orlove and Wood (1978), Michod and Hamilton (1980), Uyenoyama and Feldman (1981), Aoki (1981), Seger (1981), Pepper (2000), Goodnight (2005) and references therein. For single-locus models, relatedness measures attempt to account for the difference in the frequency of an altruistic allele in recipients versus the general population, and hence can depend on genotypic frequencies and other details of

the assumed model. See Toro et al. (1982) and Michod (1982) for a more detailed discussion of these population-genetic models. There are a very large number of models for altruism and cooperation build around Hamilton's rule, see Lehmann and Keller (2006a,b and references therein) for an overview.

Queller's Generalization of Hamilton's Rule

Equation 18.56b followed by considering fitness regressed on phenotype. This makes eminent sense, as quantitative genetics tries to work with measurable quantities, such as the effect of phenotype on fitness and the correlation between breeding and phenotypic values. However, Queller (1992b) notes that by considering the regression of fitness directly on *breeding values*, an exact expression for Hamilton's rule can be obtained. The key is that, by construction (e.g., LW Chapter 3), the residuals are uncorrelated with the predictor variables in the regression. Hence, writing fitness as a multiple regression on the breeding values of the actor (i) and recipient (j),

$$W_i = a + \beta_{W_d|G} G_i + \beta_{W_s|G} G_j + e_i \quad (18.57a)$$

we always have $\sigma(e_i, G_i) = 0$, and (assuming no transmission bias so that we can ignore the second term in 18.54), Price's equation gives

$$\begin{aligned} \Delta G &= \sigma(W_i, G_i) = \sigma(a + \beta_{W_d|G} G_i + \beta_{W_s|G} G_j + e_i, G_i) \\ &= \beta_{W_d|G} \sigma(G_i, G_i) + \beta_{W_s|G} \sigma(G_j, G_i). \end{aligned} \quad (18.57b)$$

Dividing both sides by $\sigma(G_i, G_i)$ gives Queller's generalization of Hamilton's rule as

$$\beta_{W_d|G} + \beta_{W_s|G} \frac{\sigma(G_j, G_i)}{\sigma(G_i, G_i)} > 0 \quad (18.58a)$$

Recalling the definition of a regression slope (the covariance divided by the variance of the predictor, e.g., LW Chapter 3), Queller's exact result reduces to a very satisfying form:

$$\beta_{W_d|G} + \beta_{W_s|G} \beta_{G_j|G_i} > 0 \quad (18.58b)$$

where $\beta_{W_d|G}$ is the cost, $\beta_{W_s|G}$ the benefit, and $\beta_{G_j|G_i}$ is a generalized measure of relatedness between i and j . Gardner et al. (2007) caution that while Equation 18.56b is rather general and Equation 18.58b completely general (both under the assumption of no biased transmission term), that "the cost of this generality is that it hides a lot of detail, and so naive application of Hamilton's rule may lead to mistakes." See Frank (1998) for a much more detailed discussion of potential pitfalls.

GROUP SELECTION, KIN SELECTION, AND ASSOCIATIVE EFFECTS

Kin, Group, and Multilevel Selection

There is a vast (and occasionally heated) literature in evolutionary genetics and social evolution dealing with kin selection, group selection, and inclusive fitness (selected works include Wynne-Edwards 1962, 1986; Williams 1966; Maynard Smith 1964, 1976; Wade 1978; Wilson 1983; Frank 1998). Much of the debate has revolved around the evolutionary mechanism(s) needed to account for traits that reduce the fitness of an individual but increase the fitness of a group. As we have seen, Hamilton's rule gives the condition (a sufficiently high degree of relatedness among the interacting individuals) for such an allele to spread under individual

selection. This is a **kin selection** model (Maynard-Smith 1964), where interactions among kin generate an inclusive fitness that can allow an allele potentially harmful to the individual, but helpful to a group, to spread.

An alternative school, **group selection**, was proposed by Wynne-Edwards (1962, 1963), which states that traits favoring a group spread via selection at the level of groups — those carrying alleles for beneficial group behavior tend to leave more groups than those that lack them (the roots of this idea trace back to Darwin 1871). In animal and plant breeding this is not at all controversial, as family selection (choosing only those individuals from the best families) is widely practiced (Chapter 17). In its most extreme form (e.g., Wynne-Edwards 1962, 1963), group selection is envisioned to occur through isolated demes that undergo differential extinction and propagation. Quantitative-genetic models of response to group selection due to population structure have been examined by Slaktin and Wade (1978), Slaktin (1981), Crow and Aoki (1982) and Tanaka (1996). A less extreme form is the **levels-of-selection** (or **multilevel selection**) argument, where the fitness of an individual is influenced by those individuals with which it interacts, so that fitness is a function of the collection of interacting individuals, rather than a single individual. Levels-of-selection does not require isolated units, and groups (here, simply sets of interacting individuals) can form anew each generation even in a panmictic population.

In evolutionary biology, the debate over the relative importance of kin vs. group selection at times has had the feel of a holy war. One issue against group selection were concerns about the formation and subsequent propagation of groups, as well as the generation of between-group variation on which selection can act. The heavy hand of parsimony (running the risk of getting oneself cut with Ockham's razor) is also raised against group selection — why invoke it if individual selection will do (Williams 1966)? Of course, one could argue this is entirely the wrong prior. Most biologists would not disagree with the idea that an individual's fitness is also influenced by those individuals with which it interacts, so that levels-of-selection is the more reasonable default, especially since it includes individual selection as a special case. The issue then becomes an empirical one, namely the relative fitness weights on group vs. individual components. Estimates of these level of selection components is examined in detail in Chapter 24.

Much Ado About Nothing?

The reader unfamiliar with the evolutionary literature might be a bit perplexed about this controversy, as when placed in a framework of associative effects, both group and kin selection components arise and have symmetric roles (e.g., Equation 18.29c). Indeed, a number of workers have suggested group and kin selection are simply two extremes on the continuum of potential interactions and hence are closely related (Wade 1980, Queller 1991, Lehmann et al. 2007). Bijma and Wade (2008) succinctly make the point that

the ongoing debate on equivalence of kin and levels-of-selection models is partly caused by the fact that levels-of-selection models tend to hide the relatedness component of response to selection, whereas kin selection models tend to hide the multilevel selection component of response to selection. ... the response to selection is naturally described by the combination of relatedness and the degree of multilevel selection, rather than by focusing on one or the other of the two factors.

As we will see, in the absence of associative effects influencing trait value ($\sigma_{A_s}^2 = 0$) both kin and multilevel selection are required for social selection to have a response that differs from the breeder's equation. When associative effects do influence trait values, kin and multilevel selection appear as symmetric roles, and even in their absence response can still differ from the breeder's equation. These results were first clearly stated by Bijma and Wade (2008), and we closely follow their development below.

Group and Kin Selection: Models Without Trait Associative Effects

Most models of kin and group selection assume that the trait of interest is not influenced by associative effects, so that we can decompose the phenotypic value of individual i as

$$z_i = A_i + E_i,$$

where A_i denotes i 's breeding value (we use A rather than A_d to stress that this model assumes *no* associative effects, so that no A_s terms appear). While no associative effects influence the *phenotypic value* of the trait of interest, we then assume that the fitness of an individual is influenced by the phenotypes of its neighbors. This results in the *fitness* showing direct and associative effects. This distinction between models where the trait does not show associative effects while fitnesses do versus models where *both* the trait and fitness show associative effects is subtle, but rather important, as the resulting behavior is quite different (Bijma and Wade 2008).

In particular, Bijma and Wade show that when traits lack associative effects, the response under either kin or group selection deviates from the breeder's equation only when *both* relatedness and multilevel selection occur. A slightly more general development of their result proceeds as follows. First consider a classical kin selection model, where the fitness of individual i is a function of its phenotypic value plus contributions that depend on the phenotypic values of the $n - 1$ individuals with which it interacts. Generalizing Equation 18.55b, we can express i 's fitness as the multiple regression

$$W_i = a + \beta_{W_d|z} z_i + \beta_{W_s|z} \sum_{j \neq i}^n z_j + e_i \quad (18.59a)$$

where $\beta_{W_d|z}$ is the direct effect on fitness and $\beta_{W_s|z}$ the indirect (or social) effect on i 's fitness given the phenotypes of its conspecifics. Assuming the residual e_i is uncorrelated with i 's breeding value A_i , Equation 18.56a generalizes to

$$R = \Delta A = \beta_{W_d|z} \sigma(z_i, A_i) + \beta_{W_s|z} \sum_{j \neq i}^n \sigma(z_j, A_i) \quad (18.59b)$$

As mentioned previously, we can think of $\beta_{W_d|z} = c$ as the cost and $\beta_{W_s|z} = b$ as the benefit, so that for altruistic traits $\beta_{W_d|z} < 0$ and $\beta_{W_s|z} > 0$. For the infinitesimal model under random mating,

$$\sigma(z_i, A_i) = \sigma(A_i, A_i) = \sigma_A^2, \quad \text{and} \quad \sigma(z_j, A_i) = \sigma(A_j, A_i) = r_{ij} \sigma_A^2$$

If we assume that all interacting pairs have the same relationship (so that $r_{ij} = r$), Equation 18.59b reduces to

$$R = \Delta A = \sigma_A^2 \left(\beta_{W_d|z} + r(n-1)\beta_{W_s|z} \right) \quad (18.59c)$$

Hence, the trait increases when $\beta_{W_d|z} + r(n-1)\beta_{W_s|z} > 0$ (Hamilton's rule). The multilevel selection connection that appears in kin selection models is seen by defining

$$g = \frac{\beta_{W_s|z}}{\beta_{W_d|z}} \quad (18.60a)$$

Using this definition, we can rewrite Equation 18.59a as

$$W_i = a + \beta_{W_d|z} \left(z_i + g \sum_{j \neq i}^n z_j \right) + e_i = a + \beta_{W_d|z} I_i + e_i \quad (18.60b)$$

where

$$I_i = z_i + g \sum_{j \neq i} z_j$$

is the index given by Equation 18.27a. Individual selection corresponds to $g = 0$, as $I_i = z_i$. Likewise, $g = 1$ corresponds to group selection, as

$$I_i = \sum_{j=1}^n z_j = n\bar{z}$$

so that fitness is entirely a function of group mean. As above, g need not be restricted to between zero and one. For example, negative values of g correspond to selection based on deviation within a group (see Example 18.9).

Expressed in terms of g , the expected response under kin selection (Equation 18.59c) becomes

$$\Delta A = \beta_{W_d|z} \sigma_A^2 \left(1 + (n-1)gr \right) \quad (18.61a)$$

This equation makes Bijma and Wade's main point: The response is a function of the *product* of relatedness (r) and levels-of-selection (g). If either is zero, Equation 18.62a reduces to

$$\Delta A = \beta_{W_d|z} \sigma_A^2 = \frac{\sigma(W_d, z_i)}{\sigma_z^2} \sigma_A^2 = S \frac{\sigma_A^2}{\sigma_z^2} = h^2 S \quad (18.61b)$$

Hence, for response to differ from the standard breeder's equation requires *both* relatedness ($r > 0$) and multilevel selection ($g \neq 0$). While the relatedness is obvious in kin-selection formulations, the levels-of-selection component historically has been a bit less transparent, being "hidden" in the costs and benefits, $\beta_{W_s|z}$ and $\beta_{W_d|z}$ (Bijma and Wade 2008).

Now consider the response under multilevel selection. Here, fitness is a function of both individual and group value, which is usually expressed as the components of fitness due to within-group deviation ($\Delta z_i = z_i - \bar{z}$) and the group mean ($\bar{z} = \sum z_i / n$). As above, the fitness of i can be expressed as the regression

$$W_i = a + \beta_{W|\bar{z}} \bar{z} + \beta_{W|\Delta z} \Delta z_i + e_i \quad (18.62a)$$

Individual selection occurs when $\beta_{W|\bar{z}} = \beta_{W|\Delta z} = \beta$, as Equation 18.62a reduces to $W_i = a + \beta z_i + e_i$. Again assuming that e_i is uncorrelated with A_i , Price's theorem gives the response as $\Delta A = \sigma(W_i, A_i)$, which from Equation 18.62a becomes

$$R = \Delta A = \beta_{W|\bar{z}} \sigma(\bar{z}, A_i) + \beta_{W|\Delta z} \sigma(\Delta z_i, A_i) \quad (18.62b)$$

The first covariance term is given by

$$\sigma(\bar{z}, A_i) = \frac{1}{n} \sigma \left(\sum_{j=1}^n z_j, A_i \right) = \frac{\sigma(z_i, A_i) + \sum_{j \neq i} \sigma(z_j, A_i)}{n} = r_n \sigma_A^2 \quad (18.62c)$$

where $r_n = r + (1-r)/n$ (Equation 17.7a). This is just the between-group genetic variance (Chapter 17). Likewise,

$$\sigma(\Delta z_i, A_i) = \sigma(z_i - \bar{z}, A_i) = \sigma_A^2 (1 - r_n) \quad (18.62d)$$

which is the within-group genetic variance (Chapter 17). Substitution of Equations 18.62c and d into Equation 18.62b gives the expected response to multilevel selection as

$$\Delta A = \sigma_A^2 \left(r_n \beta_{W|\bar{z}} + (1 - r_n) \beta_{W|\Delta z} \right) \quad (18.62e)$$

As noted by Wade (1980) and Cheverud (1985), total response is the sum of the between-group response, $r_n \sigma_A^2 \beta_{W|\bar{z}}$, plus the within-group response, $(1 - r_n) \sigma_A^2 \beta_{W|\Delta z}$. Relatedness enters into multilevel selection models because r influences the relative amounts of within-versus between-group variance. Increasing r increases the between-group variation $r_n \sigma_A^2$ while decreasing the within-group variation $(1 - r_n) \sigma_A^2$. With increasing relatedness, more of the response comes from between-group (as opposed to within-group) selection. In the absence of relatives within the group ($r = 0$ so that $r_n = 1/n$), Equation 18.62e becomes

$$\Delta A = \sigma_A^2 \left(\frac{1}{n} \beta_{W|\bar{z}} + \left(\frac{n-1}{n} \right) \beta_{W|\Delta z} \right) \quad (18.62f)$$

Example 18.17 shows that the term in the parentheses reduces to $\beta_{W_d|z}$, and hence the response reduces to Equation 18.61b, namely the breeder's equation. Relatedness of group members is thus required for the response under the multilevel fitness model (Equation 18.62a) to depart from the breeder's equation.

Example 18.17. Following Bijma and Wade, we can formally show the connections between the fitness models given by Equation 18.59a (kin selection) and Equation 18.62a (multilevel selection). Expanding \bar{z} and Δz_i to express them in terms of z_i and $\sum_{j \neq i} z_j$, we have that

$$\begin{aligned} \beta_{W|\bar{z}} \bar{z} + \beta_{W|\Delta z} (z_i - \bar{z}) &= \beta_{W|\Delta z} z_i + (\beta_{W|\bar{z}} - \beta_{W|\Delta z}) \bar{z} \\ &= \beta_{W|\Delta z} z_i + \frac{1}{n} \left(\beta_{W|\bar{z}} - \beta_{W|\Delta z} \right) \left(\sum_{j \neq i} z_j + z_i \right) \\ &= \frac{1}{n} \left([n-1] \beta_{W|\Delta z} + \beta_{W|\bar{z}} \right) z_i + \frac{1}{n} \left(\beta_{W|\bar{z}} - \beta_{W|\Delta z} \right) \sum_{j \neq i} z_j \end{aligned} \quad (18.63a)$$

Matching terms with Equation 18.59a, the regression slopes for fitness in these two model are related as follows:

$$\beta_{W_d|z} = \frac{\beta_{W|\bar{z}} + (n-1) \beta_{W|\Delta z}}{n} \quad \text{and} \quad \beta_{W_s|z} = \frac{\beta_{W|\bar{z}} - \beta_{W|\Delta z}}{n} \quad (18.63b)$$

Similarly, we can show that

$$\beta_{W_d|z} z_i + \beta_{W_s|z} \sum_{j \neq i} z_j = (\beta_{W_d|z} - \beta_{W_s|z}) \Delta z_i + (\beta_{W_d|z} + (n-1) \beta_{W_s|z}) \bar{z} \quad (18.64a)$$

implying

$$\beta_{W|\Delta z} = \beta_{W_d|z} - \beta_{W_s|z} \quad \text{and} \quad \beta_{W|\bar{z}} = \beta_{W_d|z} + (n-1) \beta_{W_s|z} \quad (18.64b)$$

Thus, in the absence of associative effects ($\sigma_{A_s}^2 = 0$), both models are equivalent and simply differ by shifting focus over individual versus group values.

Finally, we can use these fitness relationships to show that

$$\frac{1}{n}\beta_{W|\bar{z}} + \left(\frac{n-1}{n}\right)\beta_{W|\Delta z} = \beta_{W_d|z}$$

Using Equation 18.64b to substitute for $\beta_{W|\bar{z}}$ and $\beta_{W|\Delta z}$ gives

$$\frac{\beta_{W_d|z} + (n-1)\beta_{W_s|z}}{n} + \left(\frac{n-1}{n}\right)(\beta_{W_d|z} - \beta_{W_s|z}) = \beta_{W_d|z}$$

showing the response to multilevel selection in the absence of relatedness ($r = 0$) reduces to the breeder's equation,

$$\Delta A = \sigma_A^2 \left(\frac{1}{n}\beta_{W|\bar{z}} + \left(\frac{n-1}{n}\right)\beta_{W|\Delta z} \right) = \sigma_A^2 \beta_{W_d|z} = h^2 S$$

Hence, relatedness is required for the response under multilevel selection to deviate from the breeder's equation (Bijma and Wade 2008).

As mentioned, although these models have no associative effects when *trait* values are considered, the fitness functions (Equation 18.59a and 18.62a) generate direct and associative effects in *fitness*. It will provide very useful (especially when contrasting the above results with models that *do* allow traits to have associative effects) to consider the direct and associative components of fitness as they relate to the breeding value A_i of the focal individual. Write the index I_i as

$$I_i = z_i + g \sum_{j \neq i}^n z_j = A_i + g \sum_{j \neq i}^n A_j + e_i^*$$

where we have swept all of the non-breeding value terms into the residual e_i^* . Substitution into Equation 18.60b gives the fitness of i in terms of the breeding values A_i, A_j as

$$W_i = \beta_{d|z} A_{d,i} + g\beta_{d|z} \sum_{j \neq i}^n A_{d,j} + \epsilon_i^* = W_{d,i} + \sum_{j \neq i}^n W_{s,j} + \epsilon_i^* \quad (18.65a)$$

We have now reverted to the A_d notation for the breeding values (of direct effects), as we will shortly expand this result to allow for breeding values of associative effects (A_s). The right side of Equation 18.65a decomposes the fitness into direct and associative components, with

$$W_{d,i} = \beta_{d|z} A_{d,i}, \quad \text{and} \quad W_{s,j} = g\beta_{d|z} A_{d,j} \quad (18.65b)$$

The direct component ($W_{d,i}$) is the contribution from genes in the focal individual i towards its fitness, while the associative component ($W_{s,j}$) is the contribution from genes in j towards i 's fitness. From Price's equation, we have the response as the sum of direct and associative contributions, where

$$R_{i,d} = \sigma(W_{d,i}, A_i) = \beta_{d|z} \sigma(A_{d,i}, A_i) = \beta_{d|z} \sigma_{A_d}^2 \quad (18.65c)$$

which is the breeders equation, while any additional contribution from i due to genes in j is given by

$$R_{i,j} = \sigma(W_{s,j}, A_i) = g\beta_{d|z} \sigma(A_{d,j}, A_i) = rg\beta_{d|z} \sigma_{A_d}^2 \quad (18.65d)$$

Thus, two factors are required for genes in j to influence the response based on selecting i . First, multilevel selection ($g \neq 0$) is required in order for there to be an effect of genes in j on the *fitness* of i (here $g\beta_{d|z}A_{d,j}$). Second, for this to induce a change in i further requires that the genes in i and j are correlated ($r > 0$). Only when *both* occur is there an additional increment from the genes in j , and only in this case do we see departures from the breeder's equation. Summing over all $n - 1$ of i 's interacting neighbors recovers Equation 18.61a.

Group and Kin Selection in the Associative Effects Framework

The class of models just considered assumes that trait values are not influenced by associative effects (i.e., no $A_{s,j}$ terms appear in expression for z_i), but does allow for fitnesses to be influenced by the trait values of others (i.e., Equations 18.59a and 18.62a). Under this assumption, the response to social selection only deviates from the breeder's equation when $gr \neq 0$. When associative effects are present ($\sigma_{A_s}^2 > 0$), *both* the trait value and fitness of the focal individual may be functions of the genotypes in the group. In this case, kin and group selection (r and g) play symmetric roles, and when *either* is nonzero, response can deviate from the breeder's equation. Indeed, even when both $r = g = 0$, response can still deviate from the breeder's equation (e.g., Equation 18.11a). Hence, the addition of associative effects produces a profound change in model behavior. We first formally present the expected response to selection and then discover the source(s) of this rather different behavior in the presence of associative effects by considering the contributions to direct and associative fitnesses.

When associative effects are present (A_s terms appear), from Equation 18.1b the index I becomes

$$\begin{aligned} I_i &= z_i + g \sum_{j \neq i}^n z_j = \left(A_{d,i} + \sum_{j \neq i} A_{s,j} \right) + g \sum_{j \neq i}^n \left(A_{d,j} + \sum_{k \neq j} A_{s,k} \right) + e_i \\ &= \left(A_{d,i} + g(n-1)A_{s,i} \right) + \sum_{j \neq i} \left(A_{s,j} + g \left[A_{d,j} + (n-2)A_{s,j} \right] \right) + e_i \end{aligned} \quad (18.66)$$

When associative effects are present, Equation 18.1e shows that the expected response is the change in the total breeding value $A_T = A_d + (n-1)A_s$, where A_d and A_s are the direct and social breeding values. Applying the Price Equation (assuming that $\sigma(A_{T,i}, e_i) = 0$ and no transmission bias), we have

$$R = \Delta A_T = \sigma(W_i, A_{T,i}) = \beta_{d|z} \sigma(I_i, A_{T,i}) + \sigma(e_i, A_{T,i}) = \beta_{d|z} \sigma(I_i, A_{T,i}) \quad (18.67a)$$

Equivalently, we could have used a breeder's equation framework (Equation 18.28b) by recalling that for $W_i = \beta_{d|z} I_i + \epsilon$, the regression slope can be expressed as

$$\beta_{d|z} = \frac{\sigma(W, I)}{\sigma^2(I)} = \frac{S_I}{\sigma(I)} \frac{1}{\sigma(I)} = \frac{\bar{z}_I}{\sigma(I)},$$

recovering Equation 18.28b. Note that we have used the Robertson-Price identity, which here gives $\sigma(W, I) = S_I$. Substituting the expression for $\sigma(A_{T,i}, I_i)$ given by Equation 18.29c, the response becomes

$$R = \beta_{W_d|z} \left([g + r + (n-2)gr] \sigma_{A_T}^2 + (1-g)(1-r) [\sigma_{A_d}^2 + (n-1)\sigma_{A_d, A_s}] \right) \quad (18.67b)$$

Note the completely symmetric roles of relatedness r and levels-of-selection g in the response. The term in the second set of square brackets can be negative, resulting in R and β potentially

having opposite signs (and hence a maladaptive response). Increasing either relatedness or the amount of weight g on the other individuals within the group results in increased weight on the $\sigma_{A_T}^2$ term, which is always non-negative, increasing the chance of congruence between the signs of R and β . Finally, there is a synergistic effect between r and g in groups of size greater than two, $(n-2)gr$. Bijma and Wade note that this occurs because $n-2$ is the number of group members that two individuals have in common.

Let's examine Equation 18.67b for a couple of special cases. First (as expected) in the absence of heritable social effects ($\sigma_{A_s}^2 = 0$), then $\sigma_{A_T}^2 = \sigma_{A_d}^2$, $\sigma_{A_d, A_s} = 0$ and Equation 18.67b reduces to

$$R = \beta_{W_d|z} \sigma_{A_d}^2 [1 + (n-1)gr]$$

recovering Equation 18.61a, and showing that (in this case) *both* relatedness and multilevel selection are required for the response to deviate from the breeder's equation. Now suppose heritable social effects are present, but interacting group members are unrelated. Here $r = 0$ and Equation 18.67b reduces to

$$R = \beta_{W_d|z} \left(g\sigma_{A_T}^2 + (1-g) [\sigma_{A_d}^2 + (n-1)\sigma_{A_d, A_s}] \right) \quad (18.67c)$$

Hence, when associative effects on the phenotype occur ($\sigma_{A_s}^2 > 0$), relatedness is not required for traits with social effects to show a deviation in response from the breeder's equation. Likewise, if $g = 0$ and $r > 0$, the above result is symmetric, with g replaced by r , so that when relatedness among group members is present, multilevel selection is not needed for deviation from the breeder's equation. Finally, in the absence of both kin and group selection ($g = r = 0$), Equation 18.67b reduces to

$$R = \beta_{W_d|z} \left(\sigma_{A_d}^2 + (n-1)\sigma_{A_d, A_s} \right) = \beta_{W_d|z} \sigma_{A_d}^2 + (n-1)\beta_{W_d|z} \sigma_{A_d, A_s} \quad (18.67d)$$

where the first term in the last equality is the standard breeder's equation. Hence, the presence of associative effects, by themselves, are sufficient to produce deviations from the breeder's equation even in the absence of kin or group selection, provided the direct and social breeding values are correlated. This point was highlighted earlier in the chapter. Another point stressed earlier in the chapter is that if σ_{A_d, A_s} is sufficiently negative, the direction of response R in the trait can be different from the direct selection $\beta_{w|z}$ on the trait, producing a reversed response. Bijma and Wade (2008) take this point further, and note that (by definition) selection for an altruistic trait results in a decrease in individual fitness when the mean trait value increases. Hence, spread of an altruistic trait is an example of a reversed response, and we see that if the direct and social breeding values are sufficiently negative correlated within an individual that this can happen *even in the absence* of kin or group selection. The careful reader will immediately be concerned that this result seems at odds with Example 18.16, where relatedness was required for individual selection ($g = 0$) to spread an altruistic trait. The difference is in the models. Example 18.16 assumed associative effects only in fitness, but Equation 18.67d assumes associative effects in the *trait itself*. This means that the social breeding value in individual i influences not just the fitness of a group member j , but also j 's *trait value*. Under this setting, altruism can evolve in the absence of either group selection or relatedness.

What is the basis for these dramatic differences in behavior for models with and without associative trait effects? The key is to consider the direct and associative components of individual fitness W_i , as we did in Equation 18.65b. From Equation 18.66, these are given by

$$W_{d,i} = \beta_{d|z} (A_{d_i} + g(n-1)A_{s_i}), \quad W_{s,j} = \beta_{d|z} (A_{s_j} + g[A_{d_j} + (n-2)A_{s_j}]) \quad (18.68)$$

as obtained by Bijma and Wades (2008). These generalize Equation 18.65b, reducing to them when associative trait effects are absent ($A_s = 0$). Careful inspection of these components show the sources and targets of selection, and the implications for response. Extending upon the results presented by Equation 18.67a, from the Price equation the response is given by

$$\Delta A_T = \sigma(W_i, A_{T_i}) = \sigma(W_{d,i}, A_{T_i}) + \sum_{j \neq i} \sigma(W_{s,j}, A_{T_i}) \quad (18.69a)$$

No relatedness is required to obtain a response from $\sigma(W_{d,i}, A_{T_i})$, while i and j must be related for any terms involving $W_{s,j}$ to enter into the response. This arises because Equation 18.68 expresses $W_{d,i}$ in terms of i 's direct and associative breeding values A_{d_i} and A_{s_i} , with

$$\sigma(A_{d_i}, A_{T_i}) = \sigma(A_{d_i} + (n-1)A_{s_i}, A_{d_i}) = \sigma_{A_d}^2 + (n-1)\sigma_{A_d, A_s} \quad (18.69b)$$

and

$$\sigma(A_{s_i}, A_{T_i}) = \sigma_{A_d, A_s} + (n-1)\sigma_{A_s}^2 \quad (18.69c)$$

Conversely, $W_{s,j}$ is a function of j 's direct and associative breeding values A_{d_j} and A_{s_j} , where

$$\sigma(A_{d_j}, A_{T_i}) = r_{ij}\sigma(A_{d_i}, A_{T_i}) \quad \text{and} \quad \sigma(A_{s_j}, A_{T_i}) = r_{ij}\sigma(A_{s_i}, A_{T_i}), \quad (18.69d)$$

both of which are zero if i and j are unrelated.

Equations 18.68 and 18.69a show the roles played multilevel selection g and relatedness r . Multilevel selection determines how the associative and direct breeding values of i and j are distributed over the direct and associative components of fitness, while relatedness allows the associative component of fitness to contribute to response. For example, in the absence of multilevel selection,

$$W_{d,i} = \beta_{d|z}A_{d_i} \quad \text{and} \quad W_{s,j} = \beta_{d|z}A_{s_j}$$

and one can only utilize $\sigma_{A_s}^2$ when $W_{s,j}$ enters into the response, which requires that i and j are related. Conversely, in the absence of any relatives within a group, $W_{s,j}$ does not contribute to response, with response entirely determined by

$$W_{d,i} = \beta_{d|z}(A_{d_i} + g(n-1)A_{s_i})$$

Here a nonzero g allows for A_{s_i} to be included in $W_{d,i}$, and hence $\sigma_{A_s}^2$ is incorporated into the response (Equation 18.69c). Finally, if both r and g are zero, then only $W_{d,i} = \beta_{d|z}A_{d_i}$ enters into the response, giving (from Equation 18.69d)

$$R = \Delta A_T = \beta_{d|z}\sigma(A_{d_i}, A_{T_i}) = \beta_{d|z}(\sigma_{A_d}^2 + (n-1)\sigma_{A_d, A_s})$$

which shows a departure from the breeder's equation when direct and social breeding values are correlated. Since A_s only enters through its covariance with A_d , we are not directly accessing $\sigma_{A_s}^2$, so that no direct selection on social values occurs.

Closing Comments

In the absence of associative effects influencing the trait value of interest, both relatedness and multilevel selection are required in order for the response to deviate from the breeder's equation. When associative effects are assumed (the trait value of a group member depends in part on the genotypes of other group members) then under most settings response deviates from the breeder's equation. The only situation where this is not true is when the direct

and social breeding values are uncorrelated and there is no group or kin selection. However, simple deviation from the breeder's equation (by itself) does not imply that social (i.e., group-level) traits will evolve in a favorable direction. In many cases, the deviation is entirely due to the correlation between social and direct breeding values, which can result in a correlated response in an unfavorable direction. In order for selection to directly access social breeding values, $\sigma_{A_g}^2$ must appear in the response equation, and with associative effects this only occurs when either r or g is nonzero. Hence, in the evolution of social traits, three different components are important to consider, namely the presence or absence of (i) associative effects influencing *trait* value, (ii) kin selection ($r \neq 0$), and (iii) multilevel selection ($g \neq 0$). In the presence of associative effects, beneficial changes in the mean social value typically require either kin or multilevel selection. In the absence of associative effects both kin and multilevel selection are required.

The associative effects framework is very powerful, as it brings the full machinery of quantitative genetics to bear to the evolution of group-level traits. One immediate advantage is biological, in that quantitative trait models provide a more realistic description of complex traits (be they behavioral or morphological) than do the single-locus models upon which much of the earlier work on social selection is built. The other advantage is empirical: machinery for actually estimating the breeding values (and variances/covariances) of direct and associative effects has been developed here. Machinery to estimate components of selection at different levels (individuals vs. group) is examined in Chapter 24.

Literature Cited

- Allard, R. W., and J. Adams. 1969. The role of inter-genotypic interactions in plant breeding. *Proc. XII Int. Congress Genetics* 3: 349–370. [18]
- Aoki, K. 1981. Algebra of inclusive fitness. *Evolution* 35: 659–663. [18]
- Aoki, K. 1982. Additive polygenic formulation of Hamilton’s model of kin selection. *Heredity* 49: 163–169. [18]
- Arango, J., I. Misztal, S. Tsuruta, M. Culbertson, and W. Herring. 2005. Estimation of variance components including competitive effects of Large White growing gilts. *J. Anim. Sci.* 83: 1241–1246. [18]
- Arora, V., and P. Lahiri. 1997. On the superiority of the Bayesian method over BLUP in small area estimation problems. *Statistica Sinica* 7: 1053–1063. [18]
- Bergsma, R., E. Kanis, E. F. Knol, and P. Bijma. 2008. The contribution of social effects to heritable variation in finishing traits of domestic pigs (*Sus scrofa*). *Genetics* 178: 1559–1570. [18]
- Bijma, P. 2010a. Fisher’s fundamental theorem of inclusive fitness and the change in fitness due to natural selection when conspecifics interact. *J. Evol. Biol.* 23: 194–206. [18]
- Bijma, P. 2010b. Multilevel selection 4: modeling the relationship of indirect genetic effects and group size. *Genetics* **under review** [18]
- Bijma, P. 2010c. Estimating indirect genetic effects: precision of estimates and optimum designs. *Genetics* **under review** [18]
- Bijma, P., W. M. Muir, E. D., and J. A. M. Van Arendonk. 2007a. Multilevel selection 1: quantitative genetics of inheritance and response to selection. *Genetics* 175: 277–288. [18]
- Bijma, P., W. M. Muir, E. D., Ellen, J. B. Wolf, and J. A. M. Van Arendonk. 2007b. Multilevel selection 2: Estimating the genetic parameters determining inheritance and response to selection. *Genetics* 175: 289–299. [18]
- Bijma, P., and M. J. Wade. 2008. The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J. Evol. Biol.* 21: 1175–1188. [18]
- Boyd, R., and P. J. Richerson. 1980. Effect of phenotypic variation on kin selection. *Proc. Natl. Acad. Sci. USA* 77: 7506–7509. [18]
- Cappa, E. P., and R. J. C. Cantet. 2008. Direct and competition additive effects in tree breeding: Bayesian estimation from an individual tree mixed model. *Silvae Genetica* 57: 45–56. [17]
- Cantet, R. J. C., and E. P. Cappa. 2008. On identifiability of (co)variance components in animal models with competition effects. *J. Anim. Breed. Genet.* 125: 371–381. [18]
- Cavalli-Sforza, L. L., and M. W. Feldman. 1978. Darwinian selection and “altruism”. *Theor. Pop. Biol./* 14: 268–280. [18]
- Chen, C. Y., R. K. Johnson, S. D. Kachman, and L. D. Van Vleck. 2006. Estimation of variance components due to competition effects for selected lines of swine. *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production*, Belo Horizonte, MG, Brazil, contribution 17–03. [18]
- Chen, C. Y., R. K. Johnson, S. Newman, and L. D. Van Vleck. 2008. Estimation of genetic parameters for average daily gain using models with competition effects. *J. Anim. Sci.* 86: 2525–2530. [18]
- Chen, C. Y., R. K. Johnson, S. Newman, S. D. Kachman, and L. D. Van Vleck. 2009. Effects of social interactions on empirical responses to selection for average daily gain of boars. *J. Anim. Sci.* 87: 844–849. [18]
- Cheverud, J. M. 1984. Evolution by kin selection: A quantitative genetic model illustrated by maternal performance in mice. *Evolution* 38: 766–777. [18]
- Cheverud, J. M. 1985. A quantitative genetic model of altruistic selection. *Behav. Ecol. Sociobiol.* 16: 239–243. [18]

- Cockerham, C. C., and P. M. Burrows. 1971. Populations of interacting autogenous components. *Amer. Natl.* 105: 13–29. [18]
- Cockerham, C. C., P. M. Burrows, S. S. Young, and T. Prout. 1972. Frequency-dependent selection in randomly mating populations. *Amer. Natl.* 106: 493–515. [18]
- Craig, D. M. 1982. Group selection versus individual selection: an experimental analysis. *Evolution* 36: 271–282. [18]
- Craig, J. V., and W. M. Muir. 1996. Group selection for adaptation to multiple-hen cages: beak-related mortality, feathering, and body weight responses. *Poultry Sci.* 75: 294–302. [18]
- Crow, J. F., and K. Aoki. 1982. Group selection for a polygenic behavioral trait: a differential proliferation model. *Proc. Natl. Acad. Sci. USA* 79: 2628–2631. [18]
- Darwin, C. 1871. *The descent of man, and selection in relation to Sex*. John Murray, London. [18]
- Dawson, P. S. 1969. A conflict between darwinian fitness and population fitness in *Tribolium* “competition” experiments. *Genetics* 62: 413–419. [18]
- Dickerson, G. E. 1947. Composition of hog carcasses as influenced by heritable differences in rate and economy of gain. *Res. Bul. Iowa Agr. Exp. Sta.* 354: 489–524. [18]
- Ellen, E. D., J. Visscher, J. A. M. van Arendonk, and P. Bijma. 2008. Survival of laying hens: Genetic parameters for direct and associative effects in three purebred layer lines. *Poultry Sci.* 87: 233–239. [18]
- Ellen, E. D., W. M. Muir, F. Teuscher, and P. Bijma. 2007. Genetic improvement of traits affected by interactions among individuals: sib selection schemes. *Genetics* 176: 489–499. [18]
- Engels, W. R. 1983. Evolution of altruistic behavior by kin selection: an alternative approach. *Proc. Natl. Acad. Sci. USA* 80: 515–518. [18]
- Ewens, W. J. 1989b. An interpretation and proof of the fundamental theorem of natural selection. *Theo. Pop. Biol.* 36: 167–180. [18]
- Ewens, W. J. 1992. An optimizing principle of natural selection in evolutionary population genetics. *Theo. Pop. Biol.* 42: 333–346. [18]
- Falconer, D. S. 1965. Maternal effects and selection response. In S. J. Geerts (ed.) *Genetics today, proceedings of the XI international congress on genetics, Vol 3* pp. 763–774. Pergamon, Oxford. [18]
- Frank, S. A. 1998. *The foundations of social evolution*. Princeton University Press, Princeton. [18]
- Gallais, A. 1976. Effects of competition on means, variances and covariances in quantitative genetics with an application to general combining ability selection. *Theor. Appl. Gen.* 47: 189–195. [18]
- Gardner, A., S. A. West, and N. H. Barton. 2007. The relationship between multilocus population genetics and social evolution theory. *Amer. Natl.* 169: 207–226. [18]
- Goodnight, C. J. 1985. The influence of environmental variation on group and individual selection in a cress. *Evolution* 39: 545–558. [18]
- Goodnight, C. J. 2005. Multilevel selection: the evolution of cooperation in non-kin groups. *Popl. Ecol.* 47: 3–12. [18]
- Goodnight, C. J., and L. Stevens. 1997. Experimental studies of group selection: what do they tell us about group selection in nature? *Amer. Natl.* 150: S59–S79. [18]
- Grafen, A. 2006. Optimization of inclusive fitness. *J. Theor. Bio.* 238: 541–563. [18]
- Griffing, B. 1967. Selection in reference to biological groups. I. Individuals and group selection applied to populations of unordered groups. *Aust. J. Biol. Sci.* 20: 127–139. [18]
- Griffing, B. 1968a. Selection in reference to biological groups. II. Consequences of selection in groups of one size when evaluated in groups of a different size. *Aust. J. Biol. Sci.* 21: 1163–1170. [18]
- Griffing, B. 1968b. Selection in reference to biological groups. III. Generalized results of individuals and group selection in terms of parent-offspring covariances. *Aust. J. Biol. Sci.* 21: 1171–1178. [18]

- Griffing, B. 1969. Selection in reference to biological groups. IV. Application of selection index theory. *Aust. J. Biol. Sci.* 22: 131–142. [18]
- Griffing, B. 1976a. Selection in reference to biological groups. V. Analysis of full-sib groups. *Genetics* 82: 703–722. [18]
- Griffing, B. 1976a. Selection in reference to biological groups. VI. Use of extreme forms of non-random groups to increase selection efficiency. *Genetics* 82: 723–713. [18]
- Griffing, B. 1977. Selection for populations of interaction genotypes. In E. Pollak, O. Kempthorne, and T. B. Bailey, Jr., (eds.), *Proceedings of the international conference on quantitative genetics*, pp. 413–434. Iowa State Univ. Press, Iowa. [18]
- Hadfield, J. D., and A. J. Wilson. 2007. Multilevel Selection 3: Modelling the effects of interacting individuals as a function of group size. *Genetics* 177: 667–668. [18]
- Haldane, J. B. S. 1932. *The causes of evolution*. Longmans, London. [18]
- Haldane, J. B. S. 1955. Population genetics. *Penguin New. Bio.* 18: 34–51. [18]
- Hamilton, W. D. 1963. The evolution of altruistic behavior. *Amer. Natl.* 97: 354–356. [18]
- Hamilton, W. D. 1964a. The genetical evolution of social behavior. I. *J. Theor. Biol.* 7: 1–16. [18]
- Hamilton, W. D. 1964b. The genetical evolution of social behavior. II. *J. Theor. Biol.* 7: 17–52. [18]
- Hamilton, W. D. 1970. Selfish and spiteful behaviour in an evolutionary model. *Nature* 228: 1218–1220. [18]
- Harville, D. A. 1977. Maximum likelihood approaches to variance component estimation and to related problems. *J. Am. Stat. Assoc.* 72: 320–338. [18]
- Hsu, W. L., R. K. Johnson, and L. D. Van Vleck. 2010. Effects of pen mates on growth, backfat depth, and longissimus muscle area of swine. *J. Anim. Sci.* 88: 895–902. [18]
- Jiang, J. 1996. REML estimation: asymptotic behavior and related topics. *Ann. Stats.* 24: 255–286. [18]
- Kirkpatrick, M. and R. Lande. 1989. The evolution of maternal characters. *Evolution* 43: 485–503. [18]
- Lande, R., and M. Kirkpatrick. 1990. Selection response in traits with maternal inheritance. *Genet. Res.* 55: 189–197. [18]
- Lehmann, L., and L. Keller. 2006a. The evolution of cooperation and altruism – a general framework and a classification of models. *J. Evol. Biol.* 19: 1365–1376. [18]
- Lehmann, L., and L. Keller. 2006b. Synergy, partner choice and frequency dependence: their integration into inclusive fitness theory and their interpretation in terms of direct and indirect fitness effects. *J. Evol. Biol.* 19: 1426–1436. [18]
- Lehmann, L., L. Keller, S. West, and D. Roze. 2007. Group and kin selection: two concepts but one process. *Proc. Natl. Acad. Sci. USA* 104: 6736–6739. [18]
- Levins, R. 1970. Extinction. *Am. Math. Sco.* 2: 75–108.
- Maynard-Smith, J. 1964. Group selection and kin selection. *Nature* 201: 1145–1147. [18]
- Maynard-Smith, J. 1976. Group selection. *Quart. Rev. Biol.* 51: 277–283. [18]
- McGlothlin, J. W., and E. D. Brodie III. 2009. How to measure indirect genetic effects: the congruence of trait-based and variance-partitioning approaches. *Evolution* 63: 1785–1795. [18]
- Michod, R. E. 1982. The theory of kin selection. *Ann. Rev. Ecol. System.* 13: 23–55. [18]
- Michod, R. E., and R. Abugov. 1980. Adaptive topography in family-structured models of kin selection. *Science* 210: 667–669. [18]
- Michod, R. E., and W. D. Hamilton. 1980. Coefficients of relatedness in sociobiology. *Nature* 288: 694–697. [18]
- Moore, A. J., E. D. Brodie III, and J. B. Wolf. 1997. Interacting phenotypes and the evolutionary process. I. direct and indirect genetic effects of social interactions. *Evolution* 51: 1352–1362. [18]

- Muir, W. M. 1985. Relative efficiency of selection for performance of birds housed in colony cages based on production in single bird cages. *Poultry Sci.* 64: 2239–2447. [18]
- Muir, W. M. 1996. Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science* 75: 447–458. [18]
- Muir, W. M. 2005. Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* 170: 1247–1259. [18]
- Muir, W. M. and R. D. Howard. 1999. Possible ecological risks of transgenic organism release when transgenes affect mating success: sexual selection and the Trojan gene hypothesis. *PNAS USA* 24:13853–13856. [18]
- Muir, W. M., and A. Schinckel. 2002. Incorporation of competitive effects in breeding programs to improve productivity and animal well being. *Proc. 7th World Congr. Genet. Appl. Livestock Prod.* CM-ROM Communication No. 14-07. Montpellier, France. [18]
- Orlove, M. J., and C. L. Wood. 1978. Coefficients of relationship and coefficients of relatedness in kine selection: a covariance form for the RHO formula. *J. Theor. Biol.* 73: 679–686. [18]
- Pepper, J. W. 2000. Relatedness in trait group models of social evolution. *J. Theor. Biol.* 206: 355–368. [18]
- Price, G. R. 1970. Selection and covariance. *Nature* 227: 520–521. [18]
- Price, G. R. 1972a. Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.* 36: 129–140. [18]
- Queller, D. C. 1991. Group selection and kin selection. *Trends. Ecol. Evol.* 6: 64. [18]
- Queller, D. C. 1992a. Quantitative genetics, Inclusive fitness, and group selection. *Amer. Natl.* 139: 540–558. [18]
- Queller, D. C. 1992b. A general model for kin selection. *Evolution* 46: 376–380. [18]
- Rothenberg, T. J. 1971. Identification in parametric models. *Econometrica* 39: 577–591. [18]
- Sakai, K. 1955. Competition in plants and its relation to selection. *Cold Spring Harbor Symp. Qun. Biol.* 20: 137–157. [18]
- Slaktin, M. 1981. Population heritability. *Evolution* 35: 859–871. [18]
- Slaktin, M., and M. J. Wade. 1978. Group selection on a quantitative character. *Proc. Natl. Acad. Sci. USA* 75: 3531–3534. [18]
- Schutz, W. M., C. A. Brim, and S. A. Usanis. 1968. Inter-genotypic competition in plant populations. I. feedback system with stable equilibria in populations of autogamous homozygous lines. *Crop Sci.* 8: 61–66. [18]
- Schutz, W. M., and S. A. Usanis. 1969. Inter-genotypic competition in plant populations. II. Maintenance of allelic polymorphisms with frequency-dependent selection and mixed selfing and random mating. *Genetics* 61: 875–891. [18]
- Seger, J. 1981. Kinship and covariance. *J. Theor. Biol.* 91: 191–213. [18]
- Tanaka, Y. 1996. A quantitative genetic model of group selection. *Amer. Natl.* 148: 660–683. [18]
- Taylor, P. D., and S. A. Frank. 1996. How to make a kin selection model. *J. Theor. Bio.* 180: 27–37. [18]
- Taylor, P. D., G. Wild, and A. Gardner. 2007. Direct fitness or inclusive fitness: how shall we model kin selection? *J. Evol. Biol.* 20: 301–309. [18]
- Toro, M., R. Abugov, B. Charlesworth, and R. E. Michod. 1982. Exact versus heuristic models of kin selection. *J. Theor. Biol.* 97: 699–713. [18]
- Uyenoyama, M. K., and M. W. Feldman. 1981. On Relatedness and Adaptive Topography in Kin Selection. *Theor. Pop. Biol.* 19: 87–123. [1]
- Van Vleck, L. D., and J. P. Cassady. 2005. Unexpected estimates of variance components with a true model containing genetic competition effects. *J. Anim. Sci.* 83: 68–74. [18]

- Van Vleck, L. D., L. V. Cundiff, and R. M. Kock. 2007. Effect of competition on gain in feedlot bulls from Hereford selection lines. *J. Anim. Sci.* 85: 1625–1633. [18]
- Wade, M. J. 1976. Group selection among laboratory populations of *Tribolium*. *Proc. Natl. Acad. Sci. USA* 73: 4604–4607. [18]
- Wade, M. J. 1977. An experimental study of group selection. *Evolution* 31: 134–153. [18]
- Wade, M. J. 1978. A critical review of the models of group selection. *Quart. Rev. Bio.* 53: 101–14. [18]
- Wade, M. J. 1980. Kin selection: its components. *Science* 210: 665–667. [18]
- Wiebe, G. A., F. C. Petr, and H. Stevens. 1963. Interplant competition between barley genotypes. In D. Hanson and H. F. Robinson (eds) *Statistical genetics and plant breeding*, pp. 546–555. National Academy of Sciences, NRC No. 982. [18]
- Willham, R. 1963. The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics* 19: 18–27. [18]
- Willham, R. 1972. The role of maternal effects in animal breeding. III. Biometrical aspects of maternal effects in animals. *J. Anim. Sci.* 35: 1288–1293. [18]
- Williams, G. C. 1966. *Adaptation and natural selection*. Princeton. [18]
- Wilson, D. S. 1983. The group selection controversy: history and current status. *Ann. Rev. Ecol. System.* 14: 159–187. [18]
- Wolf, J. B. 2003. Genetic architecture and evolutionary constraint when the environment contains genetics. *Proc. Natl. Acad. Sci. USA* 100: 4655–4660. [18]
- Wolf, J. B., E. D. Brodie III, J. M. Cheverud, and A. J. Moore. 1998. Evolutionary consequences of indirect genetic effects *Trends Ecol. Evol.* 13: 64–69. [18]
- Wright, A. J. 1986. Individual and group selection with competition. *Theor. Appl. Gen.* 72: 256–263. [18]
- Wynne-Edwards, V. C. 1962. *Animal dispersion in relation to social behavior*. Oliver and Boyd, Edinburgh. [18]
- Wynne-Edwards, V. C. 1963. Intergroup selection and the evolution of social systems. *Nature* 200: 623–626. [18]
- Wynne-Edwards, V. C. 1986. *Evolution through group selection*. Blackwell, Oxford. [18]
- Yokoyama, S., and J. Felsenstein. 1978. A model of kin selection for an altruistic trait considered as a quantitative character. *Proc. Natl. Acad. USA* 75: 420–422. [18]