6

Theorems of Natural Selection: Results of Price, Fisher, and Robertson

Is there some reorientation for the expression of natural selection that may provide subtle perspective, from which we can understand our subject more deeply and analyze our problems with greater ease and greater insight? My answer is · · · that the Price equation provides that sort of reorientation. — Frank (2012)

Version 20 March 2014

One of the messages from Chapter 5 is that selection, even on just one or two loci, can have very complex dynamics. Indeed, outside of the special case of frequency-independent viability selection at a single locus, selection is not even guaranteed to increase mean fitness. Despite this, the search for general theorems (exact mathematical expressions) of selection response has motivated population and quantitative geneticists for over eighty years. By shifting attention away from the details of gene-frequency change over a large number of loci to the dynamics of a single trait, the hope was that some general statements might hold. Here we focus on three classical "theorems"— **Fisher's fundamental theorem of natural selection**, **Robertson's secondary theorem of natural selection**, and **Price's theorem** — as well as **Lush's breeder's equation**, summarized in Table 6.1. Much confusion surrounds these theorems, and their inter-relationships, in the literature, and it is ironic that attempts to bring some order to population- and quantitative-genetic theory has instead resulted in a considerable amount of historical chaos.

Table 6.1. General expressions for the response of a trait to selection, with fitness as an important special case. Here w denotes relative fitness, z the value of an arbitrary trait, A_x the breeding value of trait x, R_x the total response (change in the mean in the next generation) of trait x, δR_w the partial response in mean fitness due exclusively to allele-frequency change, S_z the within-generation change in z following selection (the selection differential), and $\overline{\delta}_z$ the expected change between the mean value of a trait in selected parents and their progeny. Expressions denoted by * are true mathematical theorems, whereas the rest are approximations.

Fisher's Fundamental Theorem		Fisher (1930)
Classical interpretation	$R_w = \sigma^2(A_w)$	
Exact version*	$\delta R_w = \sigma^2(A_w)$	Price (1972b), Ewens (1989b)
Breeder's Equation	$R_z = h_z^2 S_z$	Lush (1937)
Robertson-Price Identity*	$S_z = \sigma(w, z)$	Robertson (1966), Price (1970)
Robertson's Secondary Theorem		
1966 version	$R_z = \sigma(A_z, w)$	Robertson (1966)
1968 version	$R_z = \sigma(A_z, A_w)$	Robertson (1968)
Price's Theorem*	$R_z = \sigma(w, z) + E(w\,\overline{\delta}_z)$ $= S + E(w\,\overline{\delta}_z)$	Price (1970)

As detailed below, the classic interpretation of Fisher's theorem, along with both versions of Robertson's theorem and the breeder's equation, are all approximations (albeit often good ones), not formal mathematical theorems. In contrast, Price's general expression for *any* type of selection response *is* a formal theorem, as is the Robertson-Price identity for the within-generation response in a trait due to selection. A source of confusion is that the classical view was apparently different from Fisher's own interpretation, which *is* a formal theorem but also is not very useful. As our discussion of these various "theorems", exact and otherwise, will highlight, it is the transmission of trait values from parent to offspring that generally induces complications and makes the theoretician's job challenging. This can be seen in Table 6.1 by simply comparing the exact Robertson-Price result for within-generation change with results for cross-generational change. The surprising result is not that Fisher's and Robertson's theorems are wrong, but rather that they often are reasonable to excellent approximations for much (but not all) of the dynamics of a trait under short-term selection.

Our treatment of this rather convoluted area is structured as follows. We start with Price's theorem, which is a *completely general* description of *any* selection response under *any* model of transmission. As such, it does not rely on any explicit genetic model and thus serves as an ideal platform from which to examine the other "theorems". We next turn to Fisher's fundamental theorem, which has a rich and somewhat checkered history, in part due to Fisher not being fully explicit in stating a few of his definitions. An important corollary of Fisher's theorem is that, in the absence of mutation, selection will drive the additive variance in fitness towards zero, and we next examine some of the biological implications, and misunderstandings, of this result. We then turn to Robertson's secondary theorem which focuses on the response for any arbitrary trait, not just fitness. Finally, because the breeder's equation is the basic workhorse result for much of selection-response theory in quantitative genetics, we conclude by examining its robustness in some detail.

PRICE'S GENERAL THEOREM OF SELECTION

The thoughtful reader might ask if there is a general, assumption-free statement about selection response under *any* situation. Actually, there is, namely **Price's theorem** (Price 1970, 1972a), also referred to as the **Price equation**. Price's theorem provides a notationally elegant way to describe *any* evolutionary response. It makes *no assumptions* about the mechanism of transmission of a trait from some ancestoral category (such as an individual or group) to its descendants. As such, it works for traits transmitted by standard genetics, epigenetics, and culture, and as such has been applied to everything from the evolution of languages to community structure in ecology. There is a vast literature on its application, much of it outside of genetics and evolution. Recent reviews include Frank (1995, 1997, 1998, 2012), Rice (2004), Boyd and Richerson (2005), Okasha (2006), Gardner (2008), and Helanterä and Uller (2010).

The Life and Times of George Price

George Price was one of the most enigmatic figures in modern evolutionary biology (Frank 1995, Schwartz 2000, and Harman 2011 all review Price's life and contributions). After obtaining a Ph.D. in Chemistry from the University of Chicago, he worked on the Manhattan Project before joining IBM as an engineer. At age 44, Price quit his job and started working under Cedric Smith at University College London (from 1968 -1974). In this brief tenure, he only published four solo papers and was the co-author on two others. In his roughly 25 total pages of publications, Price made three fundamental contributions to modern evolutionary theory. In addition to Price's theorem, he introduced the power of game theory to evolutionary biology (Maynard Smith and Price 1973), and was the first to fully grasp what Fisher

had really meant by his enigmatic fundamental theorem (Price 1972b). He left academia in 1974, working as a night janitor before giving away all his worldly possessions to homeless alcoholics, dying by his own hand in 1975 as a squatter in one of the poorest areas of London.

Derivation of Price's Theorem

Price's theorem expresses the expected selection response in a trait in terms of covariances, as opposed to any explicit model of transmission. It is a *completely general description of any evolutionary response*. As succinctly stated by Rice (2004), "it is an exact characterization of a relationship that must hold between phenotype, fitness, selection, and evolution." The key to Price's equation is to first consider the effect of selection on *specific categories* (measured by how many descendants each leaves) and then consider how trait values may differ between an ancestral category and its descendants. This is an extremely subtle shift in focus, one that is easy to miss and easy to misinterpret. This shift in focus decomposes response into a selection term and a remainder term due to all other forces, such as (but not limited too) imperfect transmission.

Consider selection first. Suppose there are N categories in the population, where q_i and z_i respectively denote the frequency of category i and the mean value for the trait of interest over all members of this category. Note that z can be *any* trait measure. For example, if x denotes the value of a trait, taking $z = (x - \mu_x)^2$ or $z = (x - \mu_x)^4$ gives the response in the variance and fourth moment, respectively. Averaging over all catergories, the mean trait value becomes

$$\overline{z} = \sum_{i=1}^{N} q_i \, z_i \tag{6.1}$$

Suppose that the members of category i leave a total of n_i descendants (the absolute fitness W_i for that category). The average number of descendants per category (the mean fitness) is

$$\overline{W} = \sum_{i=1}^{N} q_i \, W_i, \tag{6.2a}$$

giving the relative fitness as $w_i = W_i/\overline{W}$. Define q_i' as the frequency of *all descendants* that have category i as their *ancestor*, where

$$q_i' = w_i q_i \tag{6.2b}$$

This subtle shift in focus is Price's key idea, as q_i' is *not* the frequency of descendants *in* category i, but rather the fraction of all descendants that are *from* ancestors in category i. The focus is entirely on the categories of ancestors, *not* on which categories the descendants are in. As an example of this shift in focus, suppose our three categories of interest are the genotypes **AA**, **Aa**, and **aa** at a diallelic locus. In a traditional population-genetics analysis (Chapter 5), we write down equations to describe how the frequency of each category changes. Price used a different focus, considering instead the frequency of the *descendants* that come from each category. Suppose category $i = 2$ corresponds to **Aa** and imagine an extreme case where only **Aa** individuals survive. Here $q_2' = 1$, as all offspring have **Aa** parents. However, in the next generation (before selection), segregation gives the genotypes **AA**, **Aa**, and **aa** at frequencies 0.25, 0.5, and 0.25, so that all three categories are present *in the offspring*, but all have only **Aa** *parents*.

Now consider the transmission phase (which more generally includes everything other than selection). Let $\overline{z_i'}$ denote the mean value of the descendants from category i, which we can decompose as

$$\overline{z_i'} = z_i + \overline{\delta}_i, \tag{6.3a}$$

namely the mean value z_i of their ancestors plus a deviation $\overline{\delta}_i$. Taking the average over all ancestral categories, the average trait value over all the descendants becomes

$$\overline{z'} = \sum_i q_i' \overline{z_i'} \tag{6.3b}$$

Recalling Equations 6.1, 6.2b, and 6.3b, the response in trait value $R_z = \Delta\overline{z} = \overline{z'} - \overline{z}$ becomes

$$\begin{aligned} R_z &= \sum_i q_i' \overline{z_i'} - \sum_i q_i z_i \\ &= \left(\sum_i q_i' z_i - \sum_i q_i z_i \right) + \left(\sum_i q_i' \overline{z_i'} - \sum_i q_i' z_i \right) \\ &= \sum_i (q_i' - q_i) z_i + \sum_i q_i' \left(\overline{z_i'} - z_i \right) \\ &= \sum_i \Delta q_i z_i + \sum_i q_i' \overline{\delta}_i \end{aligned} \tag{6.4}$$

The second line follows by adding and subtracting $\sum q_i' z_i$ and suitably gathering terms. This version of the Price equation is due to Frank (1997, 2012). The first term ($\Delta q_i z_i$) represents the change due to selection based entirely on *ancestral values*, or the **partial evolutionary change caused by natural selection** (Price 1972b; Ewens 1989b; Frank 2012). The second term ($q_i' \overline{\delta}_i$) is the part of total change caused by *imperfect transmission* of ancestral values to their descendants. Equation 6.4 is conceptually very powerful, as it decomposes the response into two separate components, one strictly based on the nature of selection (Δq_i), the other on transmission ($\overline{\delta}_i$).

Both these terms can be expressed as expectations to recover the more traditional form of Price's equation (Equation 6.6). For the first term,

$$\begin{aligned} \sum_i \Delta q_i z_i = \sum_i (w_i q_i - q_i) z_i = \sum_i w_i z_i q_i - \sum z_i q_i = E(w_i z_i) - E(z_i) \cdot 1 \\ = E(w_i z_i) - E(z_i) \cdot E(w_i) \\ = \sigma(w_i, z_i), \end{aligned} \tag{6.5a}$$

while

$$\sum_i q_i' \overline{\delta}_i = \sum_i q_i w_i \overline{\delta}_i = E(w_i \overline{\delta}_i). \tag{6.5b}$$

Substituting these into Equation 6.4 yields (Price 1970, 1972a)

$$R_z = \overline{z'} - \overline{z} = \sigma(w_i, z_i) + E(w_i \overline{\delta}_i) \tag{6.6}$$

Note that all expectations are computed with respect to the *pre-selection* frequencies of the ancestral categories (q_i).

There are several equivalent expressions for the Price equation that are useful in that they emphasize different features. For example, writing

$$E(w_i \overline{\delta}_i) = \sigma(\, w_i, \overline{\delta}_i) + E(w_i) \cdot E(\overline{\delta}_i) = \sigma(\, w_i, \overline{\delta}_i) + E(\overline{\delta}_i) \tag{6.7a}$$

and substituting this result in Equation 6.6 yields

$$R_z = \sigma(w_i, z_i) + \sigma(\, w_i, \overline{\delta}_i) + E(\overline{\delta}_i). \tag{6.7b}$$

For notational ease in the rest of the chapter, we generally denote the mean value of offspring from catergory i by $\overline{z}_i$ (in place of the more cumberson $\overline{z_i'}$), so that $z_i + \overline{\delta}_i = \overline{z}_i$. Using this to combine the first two covariances in Equation 6.7b gives

$$R_z = \sigma(w_i, \overline{z}_i) + E(\overline{\delta}_i) \tag{6.7c}$$

Equation 6.7c relates the selection response to the covariance between the fitness w_i of an ancestor and the mean value $\overline{z}_i$ of its descendants. The second term, $E(\overline{\delta}_i)$, is often thought of as the expected change in mean value from ancestor to descendant in the absence of selection. As we will see at the end of the chapter, this interpretation is not quite correct, as we can further decompose this last term into a change in the absence of selection and a further change due to selection (Equation 6.37).

Summarizing, if z_i denotes the average value in category i, which has frequency q_i before selection and q_i' following selection, and whose offspring have average value $\overline{z}_i = z_i + \overline{\delta}_i$, then equivalent forms of Price's theorem are

$$R_z = \begin{cases} \sum_i \Delta q_i z_i + \sum_i q_i' \overline{\delta}_i \\ \sigma(w_i, z_i) + E(w_i \overline{\delta}_i) \\ \sigma(w_i, z_i) + \sigma(w_i, \overline{\delta}_i) + E(\overline{\delta}_i) \\ \sigma(w_i, \overline{z}_i) + E(\overline{\delta}_i) \end{cases} \tag{6.8}$$

Note from Table 6.1 that the selection differential S can be substituted for $\sigma(w_i, z_i)$ in the middle two expressions. As is shown below, many of the standard approximations for evolutionary response (breeder's equation, Fisher's fundamental theorem, Robertson's secondary theorem) follow directly from the $\sigma(w_i, z_i)$ selection term under the assumption that the residual term $E(w_i \overline{\delta}_i)$ is zero. Hence, they fail when this term is significant.

While our discussion of Price's theorem typically assumes the ancestor is a parent or midparent and the descendants are their offspring in the next generation, it holds for *any* time interval we wish to consider and for *any* set that we wish to take as ancestors (such as a group of individuals in our discussions of kin/group selection, Chapter 22). While Price's theorem is completely general and makes absolutely no assumptions about the mechanism of transmission of trait values from ancestors to their descendants, it *does* make the assumption that all descendants have ancestors. This may seem trivial, but it is violated by migration, wherein an individual appears in the next generation from ancestors not considered. Kerr and Godfrey-Smith (2008) generalize the Price equation to accommodate this by allowing arbitrary causal connections between ancestor and descendants.

Example 6.1. Let the ancestor i be the midparent (the average value of the two parents) whose value is z_i, with the descendants followed being their offspring in the next generation. Recalling the Robertson-Price identity (Table 6.1), the Price Equation becomes

$$R_z = \sigma(w_i, z_i) + E(w_i \overline{\delta}_i) = S + E(w_i \overline{\delta}_i), \tag{6.9a}$$

showing that the between-generation response equals the within-generation change S plus the fitness-weighted change $E(w_i \overline{\delta}_i)$ due to imperfect transmission of the ancestral value to its descendants. If the average value $\overline{z}_i$ of offspring is exactly the same as the value z_i of their ancestral midparent, then $\overline{\delta}_i = 0$ for all i and the response is simply S. However, the mean value of offspring generally differs from the average value of their parents. The simplest transmission model is a linear midparent-offspring regression, $z_{ij} = \mu + b(z - \mu) + e_{ij}$. Here

z_{ij} is the trait value for the jth offspring from midparent i, giving the mean value of offspring from i as $\overline{z}_i = \mu + b(z_i - \mu) + e_i$. The expected deviation becomes

$$\overline{\delta}_i = \overline{z}_i - z_i = \mu + b(z_i - \mu) + e_i - z_i = (b-1)(z_i - \mu) + e_i \qquad (6.9b)$$

Substituting this expression for $\overline{\delta}_i$ into Equation 6.9a and recalling that $E(w_i) = 1$ gives

$$E(w_i\,\overline{\delta}_i) = (b-1)\,[\,E(w_i\,z_i) - \mu\,] + E(w_i e_i).$$

From the definition of a covariance,

$$E(w_i\,z_i) = \sigma(w_i, z_i) + E(w_i)\,E(z_i) = S + 1\cdot\mu,$$
$$E(w_i\,e_i) = \sigma(w_i, e_i) + E(w_i)\,E(e_i) = \sigma(w_i, e_i) + 1\cdot 0.$$

Putting these results together gives

$$R_z = S + E(w_i\,\overline{\delta}_i) = S + (b-1)(S + \mu - \mu) + \sigma(w_i, e_i) = bS + \sigma(w_i, e_i) \qquad (6.9c)$$

Provided that the residual e_i of the midparent-offspring regression and the fitness w_i of the midparent are uncorrelated, $R = bS$. When the midparent-offspring slope is $b = h^2$, we recover the breeder's equation. While at first blush Equation 6.9c appears to be a rather general statement about the accuracy of the breeder's equation, we made a few subtle assumptions (besides the obvious one of linearity) about the parent-offspring relationship. We discuss (and remove) these in later sections of this chapter.

Example 6.2. Consider the change in allele frequency when a single diallelic locus (alleles **A** and **a**) determines viability fitness. Assume random mating among the survivors, with p = freq(**A**), and

Genotype	Frequency (before selection)	Fitness
AA	p^2	W_{AA}
Aa	$2p(1-p)$	W_{Aa}
aa	$(1-p)^2$	W_{aa}

To apply Price's theorem, we need to specify the categories to be followed, which we take as the alleles **A** and **a**. Index these by $i = 1$ and $i = 2$ (respectively), and code their associated values as $z_1 = 1$ and $z_2 = 0$. Under this coding, $\overline{z} = 1\cdot p + 0\cdot(1-p) = p$, so that R_z represents the change in p. In the absence of mutation, transmission is perfect, as the descendant allele from an **A** allele is always **A**, giving $\overline{\delta}_i = 0$. Putting these together, Equation 6.6 becomes

$$\Delta p = R_z = \sigma(w_i, z_i) + E(w_i \cdot 0) = \sigma(w_i, z_i)$$

To show that this recovers the standard population-genetic equation for allele frequency change, note that under random mating, the fitness W_1 of an **A** allele is just its marginal fitness (Equation 5.7b), $W_1 = pW_{AA} + (1-p)W_{Aa}$. Similarly, $W_2 = pW_{Aa} + (1-p)W_{aa}$ and $E(W_i) = \overline{W} = pW_1 + (1-p)W_2$, giving

$$E(W_i \cdot z_i) = \sum_{i=1}^{2} W_i \cdot z_i \cdot \text{freq(category } i) = W_1 \cdot 1 \cdot p + W_2 \cdot 0 \cdot (1-p) = p \cdot W_1,$$

Since $w_i = W_i/\overline{W}$,

$$\Delta p = \sigma(w_i, z_i) = \frac{\sigma(W_i, z_i)}{\overline{W}} = \frac{1}{\overline{W}}\Big(E(W_i \cdot z_i) - E(W_i)E(z_i)\Big)$$
$$= \frac{1}{\overline{W}}\left(p \cdot W_1 - \overline{W} \cdot p\right) = p\,\frac{(W_1 - \overline{W})}{\overline{W}},$$

which recovers Equation 5.7c.

The Robertson-Price Identity, $S = \sigma(z, w)$

When our concern is strictly on the *within-generation* change in trait value, then $\Delta\overline{z} = \mu_z^* - \mu_z$, the difference between the fitness-weighted mean after selection (but before reproduction) and the overall mean before selection, which is the selection differential S (Chapter 13). Since the within-generation change is not influenced by cross-generation transmission, any terms involving $\overline{\delta}$ are zero, and we recover the **Robertson-Price identity**

$$S = \sigma(z, w), \tag{6.10}$$

due to Robertson (1966) and Price (1970). We derive this by another route in Chapter 13, where this identity is extensively used in selection response theory. The critical insight from this identity is that no matter how complex the relationship between z and w, the (within-generation) change in mean only depends on the covariance between fitness and phenotypic value.

Recovering the Breeder's Equation

Example 6.1 showed how the breeder's equation can be recovered by using Price's theorem. That derivation made two subtle assumptions, both dealing with the midparent-offspring regression, that need to be highlighted. For now, we continue to assume linearity. In Example 6.1, this regression was $z_{ij} = \mu + b(z - \mu) + e_{ij}$, which assumes the trait mean is the same in the offspring and parental populations in the absence of selection. More generally, $z_{ij} = \mu_o + b(z - \mu_p) + e_{ij}$, in which case

$$\begin{aligned} \delta_i = \overline{z}_{i\cdot} - z_i &= \mu_o + b(z_i - \mu_p) + e_i - z_i \\ &= (b-1)z_i - b\mu_p + \mu_o + e_i \\ &= (b-1)(z_i - \mu_p) + (\mu_o - \mu_p) + e_i \end{aligned} \tag{6.11}$$

Following the same logic as in Example 6.1 yields

$$R = b\,S + \sigma(w_i, e_i) + (\mu_o - \mu_p) \tag{6.12}$$

where the last term $(\mu_o - \mu_p)$ accounts for changes (in the absence of selection) in mean phenotype from parent to offspring, for example due to the decay of linkage disequilibrium (Chapter 16), non-random associations of environmental values (Chapter 15), or inbreeding (Chapter 23, LW Chapter 10). The second subtle assumption was that the regression slope b is the same for selected and unselected parents, a point expanded on shortly.

Now suppose the parent-offspring regression is non-linear. Assuming the simplest such departure, a quadratic, makes the key point. Taking $\mu = 0$ to simplify expressions, the mean value $\overline{z}_{i\cdot}$ of offspring from a midparent with phenotypic value z_i is given by

$$\overline{z}_{i\cdot} = a + b\,z_i + c\,z_i^2 + e_i \tag{6.13a}$$

Hence,

$$\sigma(w_i, \overline{z}_{i\cdot}) = b\,\sigma\left(w_i, z_i\right) + c\,\sigma\left(w_i, z_i^2\right) + \sigma\left(w_i, e_i\right) \tag{6.13b}$$

Assuming the residuals from the midparent-offspring regression are uncorrelated with the fitness of the midparent $\sigma\left(w_i, e_i\right) = 0$, and Equation 6.13b becomes

$$\sigma(w_i, \overline{z}_{i\cdot}) = b\,S + c\,\sigma\left(w_i, z_i^2\right) \tag{6.13c}$$

From Equation 6.7c, the response is given by this expression plus $E(\overline{\delta}_i)$. With a non-linear regression, selection on the variance (and higher-order moments) can enter into the response in the mean and the strict linear version of the breeder's equation does not hold.

A final source of error when using the breeder's equation is that the regression coefficient may change when using the selected parents. To motivate this point, we follow Frank (1997) and consider any generalized linear predictor of some value z_i from k underlying predictor variables $x_{i1}, \cdots, x_{ik}$,

$$z_i = \sum_{j=1}^{k} b_j \, x_{ij} + e_i \tag{6.14}$$

For example, an individual's phenotype can be written as the sum of their allelic average effects over all k loci plus a residual error. With a single diallelic locus, $z_i = b_1 x_{i1} + b_2 x_{i2} + e_i$, where b_j is the average effect of allele j and x_{ij} is the number of copies of allele j in individual i, so that (for a diploid) $\overline{x}_{\cdot j}/2 = p_j$, the frequency of allele j. The average of Equation 6.14 becomes

$$\overline{z} = \sum_j b_j \, \overline{x}_{\cdot j} = \sum_j 2 b_j \, p_j \tag{6.15a}$$

At some time in the future, the values of both p_j (the allele frequency) and b_j (its average effect) may have changed to new values,

$$p'_j = p_j + \Delta p_j \qquad \text{and} \qquad b'_j = b_j + \Delta b_j \tag{6.15b}$$

Writing the new mean as

$$\overline{z'} = \sum_j 2 b'_j \, p'_j, \tag{6.15c}$$

the same logic leading to Equation 6.4 allows us to decompose the total change as

$$\overline{z'} - \overline{z} = \sum_j 2 b_j \Delta p_j + \sum_j 2 p'_j \, \Delta b_j \tag{6.16}$$

There are two sources for response: a change $b_j \Delta p_j$ from the change in allele frequencies weighted by their original average effects and a change $p'_j \, \Delta b_j$ in the average effects weighted by the new allele frequencies. These two terms exactly correspond to the selection and transmission terms in Price's equation (compare with Equation 6.4). Thus, $b_j \Delta p_j$ terms correspond to partial evolutionary change caused by natural selection, while the transmission term represents the response due to changes Δb_j in the average effects themselves. These changes in turn influence the parent-offspring regression slope.

FISHER'S FUNDAMENTAL THEOREM OF NATURAL SELECTION

"The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time"

This simple statement from Fisher's (1930) book (which was dictated to his wife as he paced about their living room) has generated a tremendous amount of work, discussion, and sometimes heated arguments. Fisher claimed his result was exact, a true theorem. Historically, the **classical** (and seemingly obvious) interpretation of this quote is that the rate of increase in mean fitness equals the additive variance in fitness, $R_w = \sigma^2_A(w)$. Because variances are nonnegative, this interpretation implies that mean population fitness never decreases in a

constant environment. However, we already know from Chapter 5 that this statement is incorrect, as mean fitness can decline even under simple models of selection. As a result, the mathematician Sam Karlin referred to this interpretation as "neither fundamental nor a theorem" as it requires rather special conditions to hold, especially when multiple loci influence fitness. As discussed below, the classical interpretation of Fisher's theorem holds *exactly* only under restricted conditions, but is often a good approximate descriptor. We first review the classical interpretation, and then discuss what it appears that Fisher actually meant.

The Classical Interpretation of Fisher's Fundamental Theorem

One way to demystify the classical version of Fisher's theorem is to suppose that fitness is just a trait, and use the breeder's equation $R = h^2 S$ to predict the response from selection for that trait. Letting $z = W$, and recalling $w = W/\overline{W}$, the Robertson-Price identity gives the selection differential as

$$S_W = \sigma(z, w) = \sigma(W, w) = \frac{\sigma(W, W)}{\overline{W}} = \frac{\sigma^2(W)}{\overline{W}}. \tag{6.17a}$$

Substituting into the breeder's equation gives

$$R_W = h^2_W S_W = \frac{\sigma^2_A(W)}{\sigma^2(W)} \frac{\sigma^2(W)}{\overline{W}} = \frac{\sigma^2_A(W)}{\overline{W}}. \tag{6.17b}$$

Expressed in terms of *relative fitnesses*,

$$R_w = \Delta\overline{w} = \frac{R_W}{\overline{W}} = \frac{\sigma^2_A(W)}{\overline{W}^2} = \sigma^2_A(w) \tag{6.17c}$$

recovering the classical view of Fisher's theorem. One can also use a population-genetics framework to motivate Fisher's theorem in terms of allele-frequency change. Consider a diallelic locus with constant fitnesses under random mating. The change in mean fitness is a function of the allele-frequency change Δp,

$$\Delta\overline{W} = \overline{W}(p + \Delta p) - \overline{W}(p) \tag{6.18a}$$

If the allele-frequency change is small, a first-order Taylor series approximation gives

$$\overline{W}(p + \Delta p) \simeq \overline{W}(p) + \frac{\partial\,\overline{W}}{\partial\, p}\,\Delta p \tag{6.18b}$$

or

$$\Delta\overline{W} = \overline{W}(p + \Delta p) - \overline{W}(p) = \frac{\partial\,\overline{W}}{\partial\, p}\,\Delta p \tag{6.18c}$$

From Equation 5.1a,

$$\begin{aligned} \frac{\partial\,\overline{W}}{\partial\, p} &= 2pW_{AA} + 2(1-2p)W_{Aa} + 2(p-1)W_{aa} \\ &= 2\,[\,p(W_{AA} - W_{Aa}) + (1-p)(W_{Aa} - W_{aa})\,] = 2(\alpha_A - \alpha_a) \end{aligned} \tag{6.18d}$$

where the last equality follows from the definition of the average effects (under random mating, see LW Chapter 4) of alleles A and a on fitness, namely

$$\alpha_A = p\,W_{AA} + (1-p)\,W_{Aa} - \overline{W} \qquad \text{and} \qquad \alpha_a = p\,W_{Aa} + (1-p)\,W_{aa} - \overline{W}$$

Recall that the quantity $\alpha = \alpha_A - \alpha_a$ is the average effect of an allelic substitution (LW Equation 4.6), as the difference in the average effects of these two alleles gives the mean effect on fitness from replacing a randomly-chosen a allele with an A allele.

Applying Wright's formula (Equation 5.5) together with Equation 6.18d gives

$$\Delta p = \frac{p(1-p)}{2\overline{W}} \frac{\partial \overline{W}}{\partial p} = \frac{p(1-p)}{2\overline{W}} (2\alpha) \tag{6.18e}$$

Substitution into Equation 6.18c gives

$$\Delta \overline{W} \simeq \frac{p(1-p)(2\alpha)^2}{2\overline{W}} = \frac{\sigma_A^2(W)}{\overline{W}} \tag{6.18e}$$

The last step following from the fact that the additive genetic variance is related to α by $\sigma_A^2 = 2p(1-p)\alpha^2$ (LW Equation 4.12a). Thus, under this approximation of small allele-frequency change (so that terms of order $(\Delta p)^2$ can be ignored), the change in mean fitness is indeed proportional to the additive genetic variance in fitness.

Example 6.3. The accuracy of the above first-order Taylor series approximation was examined by Li (1967). Because $\overline{W}$ is a quadratic polynomial of p, the second-order Taylor series is exact,

$$\Delta \overline{W} = \frac{\partial \overline{W}}{\partial p} \Delta p + \frac{1}{2} \frac{\partial^2 \overline{W}}{\partial p^2} \Delta p^2$$

As just shown, the first term recovers Fisher's theorem, while the second term is the error resulting from this approximation. Taking the derivative of Equation 6.18d gives

$$\frac{\partial^2 \overline{W}}{\partial p^2} = 2 \left(W_{AA} - 2W_{Aa} + W_{aa} \right)$$

Recalling Equation 6.18e, the residual term becomes

$$\frac{1}{2} \frac{\partial^2 \overline{W}}{\partial p^2} \Delta p^2 = \left(W_{AA} - 2W_{Aa} + W_{aa} \right) \left(\frac{p(1-p)}{\overline{W}} \alpha \right)^2$$

Thus, if fitnesses are additive ($[W_{AA} + W_{aa}]/2 = W_{Aa}$), the second term is zero, and Fisher's theorem holds. However, when dominance in fitness is present, Fisher's theorem fails even for a single locus under random mating.

As this example highlights, Equation 6.18e, and thus the classical view of Fisher's theorem, is only approximate. Under what conditions does this interpretation actually hold? While it is correct for multiple additive loci (i.e., no dominance nor epistasis) under both random and nonrandom mating (Kempthorne 1957; Ewens 1969), it is generally compromised by nonrandom mating and departures from additivity (such as dominance or epistasis). Even when the theorem does not hold exactly, how good an approximation is it? Nagylaki (1976, 1977a,b, 1991, 1992b, 1993) has examined ever more general models of fitness when selection is weak (the fitness of any genotype can be expressed as $1 + as$ with s small and $|\, a \,| \ll 1$) and mating is random. Selection is further assumed to be much weaker than the recombination frequency c_{min} for the closest pair of loci ($s \ll c_{min}$). Under these conditions,

Nagylaki shows that the evolution of mean fitness falls into three distinct stages. During the first phase (roughly $t < 2 \ln s / \ln(1 - c_{min})$ generations), the effects of any initial disequilibrium are moderate, first by reaching a stage where the population evolves approximately as if it were in linkage equilibrium and then by reaching a stage where the linkage disequilibrium remains relatively constant. When this latter stage is reached, we enter the central phase, where the change in mean fitness is

$$\Delta\overline{W} = \frac{\sigma_A^2(W)}{\overline{W}} + O(s^3) \tag{6.19}$$

where $O(s^3)$ means that terms on the order of s^3 have been ignored. Because additive genetic variance is expected to be of order s^2, Fisher's theorem is expected to hold to a good approximation during this period. The central phase of evolution lasts roughly $1/s$ generations. However, as gametic frequencies approach their equilibrium values, we reach the third phase, where additive variance in fitness can be much less than order s^2, in which case the error terms of order s^3 can be important. During the first and third phases, mean fitness can decrease, but the fundamental theorem holds during the central phase of evolution. Biologically, we expect the bulk of evolution to occur during this middle phase and hence Fisher's theorem approximately holds over the major part of evolutionary change. While Nagylaki's results are weak-selection approximations, we often expect weak selection to be the norm for quantitative traits, as even strong selection on a *trait* translates into weak selection on the *underlying loci* if each has a small effect (Chapter 5).

Finally, Price's theorem can be used to make an exact statement about the evolution of mean fitness. Let z be the breeding value A_W of fitness, so that $\Delta\overline{z}$ denotes the between-generation change in the mean breeding value ($\Delta\overline{A}_W$) for fitness. In particular, letting $z_i = A_i$ denote the breeding value for the ith midparent, the mean breeding value in their offspring becomes

$$\overline{z}_{i\cdot} = A_i + \overline{\delta}_i \tag{6.20a}$$

The phenotypic value of fitness for this midparent (the average of the two parental fitnesses) can be written as $W_i = A_i + \epsilon_i$, where $\sigma(A_i, \epsilon_i) = 0$ by construction. Substituting these results into Equation 6.7c gives the between-generation change in the mean breeding value ($\Delta\overline{A}_W$) for fitness as

$$\begin{aligned}
\Delta\overline{A}_W &= \sigma(w_i, A_i + \overline{\delta}_i) + E(\overline{\delta}_i) \\
&= \sigma(w_i, A_i) + \sigma(w_i, \overline{\delta}_i) + E(\overline{\delta}_i) \\
&= \frac{1}{\overline{W}}\,\sigma(A_i + \epsilon_i, A_i) + \sigma(w_i, \overline{\delta}_i) + E(\overline{\delta}_i) \\
&= \frac{\sigma_A^2(W)}{\overline{W}} + \sigma(w_i, \overline{\delta}_i) + E(\overline{\delta}_i)
\end{aligned} \tag{6.20b}$$

When $E(\overline{\delta}_i) = \sigma(w_i, \overline{\delta}_i) = 0$, we recover the classic version of the fundamental theorem. If the mean breeding value for offspring is the average of their parent's breeding values, then $\overline{\delta}_i = 0$ and these conditions hold. Even if not *exactly* true, often $\overline{\delta}_i$ is very close to zero and the leading term (and hence Fisher's theorem) dominates.

What Did Fisher Really Mean?

Fisher warned that his theorem "requires that the terms employed should be used strictly as defined", and much of the confusion stems from what Fisher meant by "fitness". Price (1972b) and Ewens (1989b, 1992) have argued that Fisher's theorem is always true, because Fisher had a very narrow interpretation of the change in mean fitness (also see Edwards 1990,

1994; Frank 1995; Lessard and Castilloux 1995; Lessard 1997; Plutynski 2006). They argue that Fisher, rather than considering the *total* rate of change in fitness, was instead concerned only with the *partial* rate of change, that due to changes in allele frequency without considering any corresponding changes in the average excesses/effects of these alleles.

Placed in the framework of Price's theorem, this partial increase interpretation becomes clear. From Equation 6.16, the total response in fitness can be decomposed into two components,

$$R_w = \sum_j 2\alpha_j \Delta p_j + \sum_j 2p'_j \, \Delta\alpha_j \tag{6.21a}$$

where α_j is the average effect of an allele on fitness. Recalling Equation 6.5a, the first sum is just $\sigma(w_i, A_i) = \sigma(A_i + \epsilon_i, A_i) = \sigma^2_A(w)$, and is given by the pre-selection values of the average effects weighted by the change in allele frequencies, yielding

$$R_w = \sigma^2_A(w) + \sum_j 2p'_j \, \Delta\alpha_j \tag{6.21b}$$

Price (1972b) stated that Fisher was interested in the change in fitness due *solely* to changes in allele frequencies, with everything else (such as average effects) held constant (also see Ewens 1989b). Thus, the exact version of Fisher's theorem simply concerns the partial response due to the first sum (the partial evolutionary response caused by natural selection), as Fisher essentially regarded the second term as a change in the "environment" within which alleles find themselves after selection. As noted by Price, Fisher had a very broad interpretation of "environment", referring to both physical and genetic backgrounds. Price claimed that Fisher

> "regarded the natural selection effect on fitness as being limited to the additive or linear effects of changes in gene (allele) frequencies, while everything else — dominance, epistasis, population pressure, climate, and interactions with other species — he regarded as a matter of the environment."

A nice discussion of this point is given by Frank and Slatkin (1992), who point out that the change in mean fitness over a generation is also influenced by the change in "environment", E. Specifically,

$$\Delta\overline{W} = \left(\overline{W}' \,\middle|\, E' \right) - \left(\overline{W} \,\middle|\, E \right) \tag{6.22a}$$

where the prime denotes the fitness/environment in the next generation. We can partition the contributions from the change in fitness and the change in the environment by writing

$$\Delta\overline{W} = \left[\left(\overline{W}' \,\middle|\, E \right) - \left(\overline{W} \,\middle|\, E \right) \right] + \left[\left(\overline{W}' \,\middle|\, E' \right) - \left(\overline{W}' \,\middle|\, E \right) \right] \tag{6.22b}$$

where the first term in brackets represents the change in mean fitness under the initial "environment", while the second represents the change in mean fitness due to changes in the environmental conditions. This decomposition is simply another way of stating Equation 6.21a. The exact version of Fisher's theorem relates solely to changes in the first component, $(\overline{W}' \,|\, E) - (\overline{W} \,|\, E)$, which he called the change in fitness due to natural selection. In considering this exact version of Fisher's theorem, Ewens (1994) states

> "I believe that the often-made statement that the theorem concerns changes in mean fitness, assumes random-mating populations, is an approximation, and is not correct in the multi-locus setting, embodies four errors. The theorem relates the so-called partial increase in mean fitness, makes no assumption about random mating, is an exact statement containing no approximation, and finally is correct (as a theorem) no matter how many loci are involved."

Nagylaki (1993) suggested that the statement $\Delta\overline{W} = \sigma^2_A(W)/\overline{W}$ be referred to as the **asymptotic fundamental theorem of natural selection**, while Fisher's more narrow (and correct)

interpretation based on partial change be referred to as the **Fisher-Price-Ewens theorem of natural selection**. A clear distinction between these two very different interpetations seems quite reasonable given their considerable past history of confusion. Warren Ewens (personal communication) said it best by noting that

> "one should think of two totally different results, holding under totally different sets of circumstances, not intersecting with each other much, and which should not be put under the same umbrella."

In summary, the bold, sweeping classical interpretation of Fisher's statement is replaced in the exact version by a standard outcome of mathematics — a result that is absolutely correct, but not really useful. While the nature of the partial response is certainly elegant, our interest is in the total response.

IMPLICATIONS OF FISHER'S THEOREM FOR TRAIT VARIATION

While the classical view of Fisher's theorem does not generally hold, and the exact version has rather limited utility, Fisher's theorem has still had an enormous impact on how quantitative geneticists view trait variation. This is because an important corollary holds under very general conditions (Kimura 1965a; Nagylaki 1976, 1977b; Ewens 1976; Ewens and Thompson 1977; Charlesworth 1987): in the absence of new variation from mutation or other sources such as migration, ***selection is expected to eventually remove all additive genetic variation in fitness***. This can be seen immediately for a single locus by considering Equation 5.8b — if the population is at equilibrium, all average excesses are zero, as all segregating alleles have the same marginal fitness and hence $\sigma^2(A_W) = 0$, so that no additive variation *in fitness* is present (Fisher 1941).

This corollary makes the general prediction that characters strongly genetically correlated with fitness should show reduced additive variance relative to characters less well correlated with fitness (Robertson 1955), reflecting the removal of additive variance by selection (which may be partly countered by new mutational input). As we now review, there is indeed a loose trend for traits correlated with fitness to show reduced heritability. However, a closer look shows that the additive variance is often *greater* for traits correlated with fitness, but that this increase is overwhelmed by an increased residual variance, resulting in a lower heritability.

Example 6.4. Consider a locus with two alleles (A_1 and A_2) and overdominance in fitness,

$$W_{11} = 1 \quad W_{12} = 1 + s \quad W_{22} = 1$$

Letting $p = \text{freq}(A_1)$, from our above definition of the average effect of a substitution under random mating,

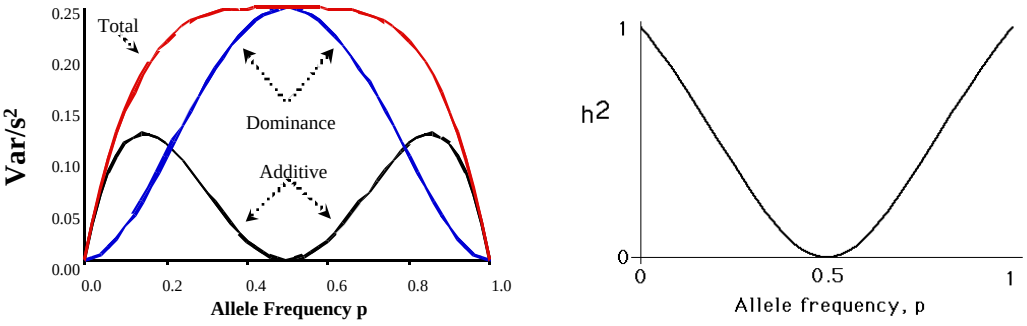
$$\alpha = \alpha_A - \alpha_a = p[\,1 - (1+s)\,] + (1-p)[\,(1+s) - 1)\,] = s(1-2p)$$

giving the additive variance in fitness as

$$\sigma_A^2(W) = 2p(1-p)\alpha^2 = 2p(1-p)s^2(1-2p)^2$$

The dominance variance is easily computed from LW Equation 4.12b to be $\sigma_D^2(W) = [2p(1-p)s]^2$. As plotted below, these variances change dramatically with p. The maximum genetic variance in fitness occurs at $p = 1/2$, but none of this variance is additive, and heritability in fitness is zero. Note that $\Delta p = 0$ when $p = 1/2$, and at this frequency $\sigma_A^2(W) = 0$,

as the corollary of Fisher's theorem predicts. Even though *total* genetic variation in fitness is maximized at $p = 1/2$, no change in $\overline{W}$ occurs as the *additive* genetic variance is zero at this frequency. For simplicity, the plot of h^2 for fitness assumes no environmental effects, so that $h^2 = \sigma_A^2/\sigma_G^2 = \sigma_A^2/(\sigma_A^2 + \sigma_D^2)$.



Even if Fisher's theorem holds exactly, it's implication for character evolution can often be misinterpreted. Suppose that locus A in this example completely determines a character under stabilizing selection. Let the genotypes AA, Aa, and aa have discrete phenotypic values of $z = -1$, 0, and 1, respectively (so that this locus is strictly additive) with respect to the trait, and let the fitness function be $W(z) = 1 - sz^2$. If we assume no environmental variance, this generates very nearly the same fitnesses for each genotype assumed above, as the fitnesses can be normalized as $1 : (1 - s)^{-1} : 1$, where $(1 - s)^{-1} \simeq 1 + s$ for small s. The additive genetic variance for the *trait* z is maximized at $p = 1/2$, precisely the allele frequency at which the additive genetic variance in *fitness* $\sigma_A^2(W) = 0$. This difference emphasizes that Fisher's theorem concerns additive genetic variance in *fitness*, not in the *character*. In this example, mapping the phenotypic character value z into the trait fitness w takes a character that is completely additive and introduces overdominance in fitness.

Traits More Highly Correlated With Fitness Have Lower Heritabilties

The corollary to Fisher's thereom, that additive variance in fitness is driven towards zero by selection, suggests that traits correlated with fitness are expected to have reduced levels of additive variance relative to characters under less direct selection. How well does this prediction hold up? Many authors have noticed that characters expected to be under selection (e.g., life-history traits, such as clutch size) tend, on average, to have lower heritabilities than morphological characters measured in the same population/species (reviewed by Robertson 1955; Roff and Mousseau 1987; Mousseau and Roff 1987; Charlesworth 1987, also see LW Figure 7.10). However, some notable exceptions are also apparent (Charlesworth 1987).

The difficulty with these general surveys is knowing whether a character is, indeed, highly *genetically* correlated with lifetime fitness. Clutch size, for example, would seem to be highly correlated with total fitness, but if birds with large clutch sizes have poorer survivorship, the correlation with lifetime fitness may be weak. Negative genetic correlations between components of fitness allow significant additive variance in each component at equilibrium, even when additive variance in *total* fitness is zero (Robertson 1955; Rose 1982).

Estimates of lifetime fitness in natural populations and their correlation with components of fitness (such as clutch size) are rare. One example is that of Gustasfsson (1986, also see Merilä and Sheldon 2000), who was able to measure lifetime reproductive success, as well as the heritabilities of fitness and other characters, in a closed natural population of collared flycatcher birds (*Ficedula albicollis*) in the Baltic Sea. Lifetime reproductive success had an

estimated heritability not significantly different from zero, as expected from the corollary to Fisher's theorem. Clutch size had a rather high heritability, 0.32 ± 0.15, but the estimated phenotypic correlation between clutch size and total fitness was very low, $r^2 = 0.03$. In general, as the phenotypic correlation between a trait and fitness decreased, its heritability increased (Figure 6.1). McCleery et al. (2004) also found a negative relationship between trait heritability and trait-fitness phenotypic correlation in an English population of Great tits (*Parus major*) followed for almost 40 years. Similar finding were seen by Teplitsky et al. (2009) in red-billed gulls (*Larus novaehollandiae*). Conversely, Schwaegerle and Levin (1991) found no significant association between the heritiability of a character and its phenotypic correlation to fruit production (chosen as one measure of total fitness) in a wild population of the plant *Phlox dummondii* (Figure 6.1). While the evidence is mixed, these studies suggest a mild trend for characters phenotypically correlated with fitness to have reduced heritabilities relative to other characters. One important caveat is that this association is based on phenotypic, rather than genetic, correlations with fitness.

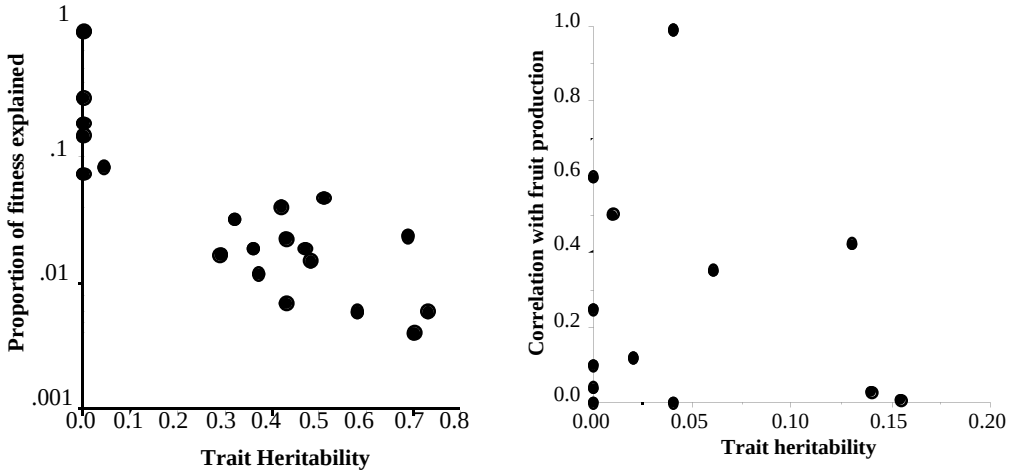


Figure 6.1. Two studies examining the association between a character's heritability and its phenotypic correlation with total fitness. **Left**: Gustafsson's (1986) work on the collared flycatcher *Ficedula albicollis* on the island of Gotland in the Baltic Sea. The percent of total fitness explained by a character (measured by r^2, the squared phenotypic correlation between the character and lifetime fitness) is plotted against the heritability of the character. **Right**: Schwaegerle and Levin's (1991) study of Texas populations of Phlox (*Phlox drummondii*). Here fruit production is used as a measure of total fitness, with the phenotypic correlation of a trait with fruit production plotted against its heritability.

Based on these observations, if a trait is known to be under selection, one might to be tempted to assume it is still far from its (phenotypic) equilibrium value if it shows a modest to large heritability. This is false. As Example 6.4 highlighted, a trait under selection can still have a high *trait* additive variance even when the additive variance in *fitness* is near zero. This can happen if there is a nonlinear transformation of the trait value (z) into fitness (w), such as occurs with stabilizing selection. More generally, Price and Schluter (1991) note that even with a simple linear relationship between a trait and fitness, one can easily have a modest heritability for the trait but a very low heritability for fitness.

The following simple model makes most of their main points. Assume fitness is entirely determined by a metric character, with fitness a linear function of the phenotypic value z,

$W(z) = a + \beta\, z + e$, the expected fitness for an individual of that phenotype ($a + \beta\, z$) plus a residual deviation e, giving the total variance in fitness as $\sigma^2[\,W(z)\,] = \beta^2\sigma_z^2 + \sigma_e^2$. Writing $z = A+E$, the additive genetic value A plus all other sources of variance (environmental and genetic), the additive variance in fitness is $\beta^2\sigma_A^2$. The heritability of fitness can be expressed in terms of the variance components for z as follows:

$$h_z^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_E^2} > h_W^2 = \frac{\beta^2\,\sigma_A^2}{\beta^2\,(\sigma_A^2 + \sigma_E^2) + \sigma_e^2} = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_E^2 + \sigma_e^2/\beta^2} \tag{6.23}$$

Thus, even when fitness is entirely determined by a single character, the heritability of fitness is less than the heritability of the character under selection, due to the residual deviation e in mapping from z to W. If the heritability of *fitness* is found to be close to zero in this case, there still could be a significant heritability in the actual *character* under selection and hence the population could still be far from a selection equilibrium for the trait value, as its mean can continue, albeit slowly, to change over time.

Traits Correlated with Fitness have Higher Levels of Both Additive and Residual Variance

One consequence of the argument made by Price and Schluter is that traits more closely associated with fitness may also have higher residual variances than traits less closely associated with fitness. While a reduced h^2 value is often interpreted as resulting from a decrease in the additive variance, it can also result from an increase in the residual variance. Thus, simple comparison of heritabilities can be misleading (Houle 1992), and a more careful examination is required to determine the cause of reduced h^2 values for fitness-related traits. If one is to compare additive variances directly across traits, a standardized measure is required. A common approach in statistics to assess relative variability is to compare coefficients of variation σ/μ, where μ is the trait mean, leading Houle (1992; also Charlesworth 1984) to suggest that the **coefficient of additive genetic variance**, $CV_A = \sigma_A/\mu$ is the appropriate scale-free measure for comparing the amount of additive genetic variation across traits. To distinguish this measure from the heritability, Houle coined the term **evolvability** for CV_A. As the representative sample of *Drosophila* traits in Table 6.2 illustrates, traits with low h^2 values can have very high CV_A values. Indeed, in a recent meta-analysis, Hansen et al. (2011) found essentially no correlation between heritability and evolvability.

Table 6.2. Heritabilities and coefficients of additive genetic (CV_A) and residual (CV_R) variation for representative traits in *Drosophilia melanogaster*. Both CV_A and CV_R values are scaled by 100. Here n is the number of studies and the median estimates are reported. After Houle (1992).

Trait	n	h^2	CV_A	CV_R
Sternopleural bristles	21	0.44	7.97	8.39
Wing length	31	0.36	2.09	1.56
Fecundity	12	0.06	39.02	11.90
Longevity	7	0.11	27.73	9.89

Surprisingly, Houle found that a survey of over 800 estimates for CV_A from a variety of traits revealed that characters assumed to be closely related to fitness (such as life-history traits) have higher evolvabilities (larger CV_A values) than do traits more loosely associated with fitness. The pattern of heritabilities decreasing with their correlation with fitness is thus not due to proportionately smaller additive variances, but rather proportionately larger residual (nonadditive plus environmental) variances, quantified by CV_R, the coefficient of residual variation.

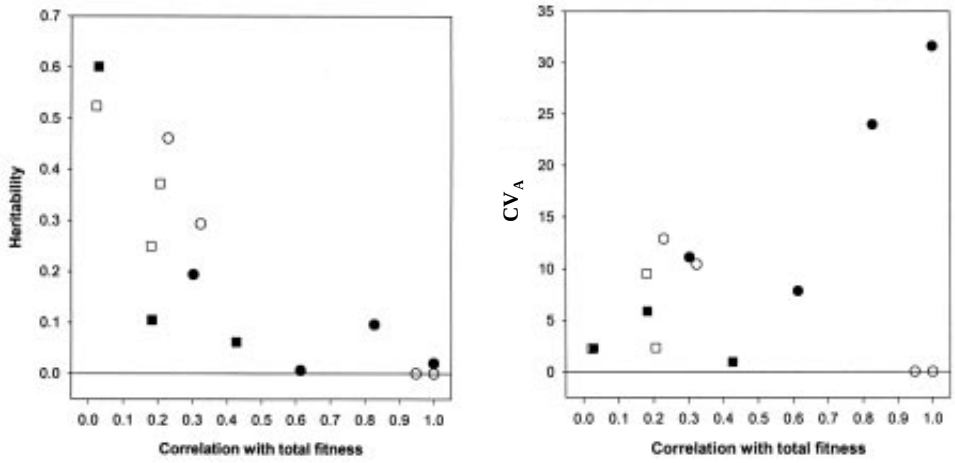


Figure 6.2. Kruuk et al.'s (2000) study of life-history and morphological traits in the Scottish red deer (*Cervus elaphus*). Circles denote life-history traits, squares morphological traits. Filled symbols are for males, open for females. **Left**: The heritability of a trait is negatively associated with the correlation of that trait with fitness. **Right**: The coefficient of additive genetic variation (the square root of the additive genetic variance of a trait divided by the trait mean, $CV_A = 100\ \sigma_A/\mu$) is positively associated with fitness in males, and negatively associated with fitness in females.

A study by Kruuk et al. (2000) on a Scottish red deer (*Cervus elaphus*) offers some additional insight. The authors estimated variance components (for additive genetic, maternal, and residual effects) from pedigree data for this wild population (on the Isle of Rum in Scotland), for five life-history and three morphological traits in addition to lifetime fitness. As shown in Figure 6.2, they also found trait heritabilities to be negatively correlated with fitness, whereas the coefficient of additive genetic variance CV_A, was *positively* correlated with fitness in males (but negatively correlated in females). Moreover, CV_A values were higher for life-history traits than for morphological traits, and the coefficient of residual variation CV_R was also positively correlated with fitness. Similar patterns of both CV_A and CV_R being positively correlated with fitness have been seen in the seed beetle *Callosobruchus maculatus* (Messina 1993), in an Alberta population of bighorn sheep (Coltman et al. 2006), in the perennial herb *Ipomopsis aggregata* (Campbell 1997), and in natural bird populations (collared flycatchers by Merilä and Sheldon 2000; great tits by McCleery et al 2004). For these studies, high residual variance, not low σ_A^2, accounts for the observed lower h^2 values for traits related to fitness. Conversely, while Teplitsky et al. (2009) also observed lower heritabilities for fitness-related traits in red-billed gulls, both higher residual variance *and* lower additive variance accounted for their trend.

What accounts for the higher additive variance in traits associated with fitness? The prediction of lower additive variance is based on the idea that σ_A^2 is removed by selection, but this is countered by new mutational input. If all traits have similar mutational variances, a faster removal of σ_A^2 by selection for traits more closely related to fitness would lead to lower equilibrium levels of additive variance. However, it appears that traits more closely associated with fitness actually have *higher* mutational variances (Houle et al. 1996), most likely do to a larger number of loci that influence fitness (Houle 1992; Houle et al. 1996; Merilä and Sheldon 1999).

While the corollary of Fisher's theorem suggests that additive genetic variance in fitness is driven to zero, mutational input counters this decline, leaving some nonzero amount even in the face of strong selection. Estimates of the actual amount of additive variance in fitness

in natural populations are very rare. Fowler et al. (1997) and Gardner et al. (2005) found significant additive variation for fitness in a laboratory population of *Drosophila* selected for domestication for close to 500 generations. Four studies have examined the additive variation associated with lifetime reproductive success in natural populations of vertebrates with extensive known pedigrees. Kruuk et al. (2000) found no evidence for a significant heritability in red deer (estimates of σ^2_A under a REML analysis were zero). McCleery et al. (2004) found positive, but not significant, estimates in great tits. Merilä and Sheldon (2000) found a significant additive variance for females and a positive (but not significant) variance for males in collared flycatchers. Finally, Teplitsky et al. (2009) found positive (but not significant) variance in females and no variance in males for red-billed gulls. One issue in all of these studies is the expected low power to detect small amounts of variances, so negative results should be viewed cautiously.

Non-Additive Genetic Variances for Traits Under Selection

As selection drives the additive variance in fitness to near zero, any remaining genetic variance is expected to be increasingly composed of non-additive terms. As Example 6.4 highlights, this non-additive variance can be considerable. Thus, characters more closely associated with fitness are expected have a higher fraction of non-additive variance. Suggestions of this trend can be seen for the results of chromosome substitution analysis (Example 6.5), which tend to show epistatic interactions for life-history characters but not for general morphological characters (also see LW Table 5.1).

Crnokrak and Roff (1995) examined roughly 340 estimates of dominance variance in both life history and morphological traits from 17 wild and 21 domestic species. In the wild species, traits assumed more closely connected with fitness (life-history traits) showed significantly higher dominance (measured as a ratio of estimated dominance to estimated total variance) than did morphological traits. In domesticated species, however, there were no significant differences in dominance between life-history and morphological traits. The presumption is that many of the morphological traits examined in the domesticated species were themselves the result of strong recent selection during domestication. This is supported by the observation that morphological traits in domesticated species showed significantly higher dominance than morphological traits in wild species. While certainly not conclusive, these results are consistent with the prediction of higher dominance genetic variance in traits more closely associated with fitness.

Roff and Emerson (2006) present a somewhat complementary analysis, using ninety estimates for life-history traits and over 140 estimates for morphological traits from line-cross data. Recall (LW Chapter 9) that line-cross analysis examines the components (additive, dominant, ect.) of the among-line variance, rather than the variance segregating in any particular population. It is by no means clear if additive variance being driven to zero by selection translates into significant non-additive components contributing to differences between line means. This caveat aside, Roff and Emerson found that the magnitude of dominance (relative to additive) effects in line differences was much greater for life-history traits. Further, epistatic effects were more often detected for life-history traits, and the ratio of total nonadditive effects (dominance plus epistasis) relative to additive effects for life-history traits was roughly double that for morphological traits. Finally, DeRose and Roff (1999) showed that (in animals) inbreeding depression is greater for life-history than morphological traits, indicating higher amounts of directional dominance among segregating alleles for life-history traits (also see LW Chapter 10).

As we have seen, there is an increase in residual variation for traits associated with fitness. What accounts for this? One source, as suggested by Price and Schluter (1991), is higher environmental variance associated with fitness. As we have just seen, a second source is an increase in nonadditive variance. While both factors likely play a role, their

relative importance is unknown (Merilä and Sheldon 1999). Due to the difficulty of estimating nonadditive genetic variance components without special mating designs, resolution of this question for natural populations is likely to prove quite difficult.

Example 6.5. Mackay (1985) examined total fitness (measured by competition against a marked balancer stock) of 41 third chromosomes extracted from a natural population of *Drosophila melanogaster*. Using these chromosomes, lines with an otherwise common background could be made homozygous for a particular extracted third chromosome, and likewise the performance of that chromosome as a heterozygote could also be assayed. If there is significant additive variance in fitness, a correlation between homozygote and heterozygote fitness is expected. Such a correlation was found for viability, suggesting some additive genetic variance in this character. However, when total fitness was examined, no correlation was found, suggesting no significant additive variation in total fitness. Mackay observed strong inbreeding depression, consistent with the total genetic variation in fitness in her study (variation in the performance of different third chromosome lines) being caused by segregation of rare deleterious recessive alleles (see LW Chapter 10). By contrast, in a very similar experiment using segregating third chromosomes within a population selected for domestication, Fowler et al. (1997) and Gardner et al. (2005) found high homozygote-heterozygote correlations, and thus significant additive variance in fitness associated with different third chromosomes.

ROBERTSON'S SECONDARY THEOREM OF NATURAL SELECTION

Like the fundamental theorem, there is confusion in the literature as to the exact form of Robertson's secondary theorem. In two separate papers, Robertson (1966, 1968) suggested that the expected change in any trait is a function of the covariance between the breeding value for that trait and relative fitness. Robertson (1968) called this relationship the **secondary theorem of natural selection**. While these two papers are usually jointly cited as the source for this theorem, a more careful reading shows Robertson proposed two *different* formulae, based on whether the covariance with trait breeding value A_z involves relative fitness w or the breeding value A_w of relative fitness.

1968 Version: $R_z = \sigma_A(z, w)$

We consider Robertson's (1968) version first, as it is widely used in the population-genetics literature (e.g., Crow and Nagylaki 1976; Nagylaki 1992b). Robertson's paper does not contain either a proof or a formal expression, just the clear statement that "The secondary theorem of natural selection states that the change in any character produced by a selection process is equal to the additive covariance between fitness and the character itself". In equation form,

$$R_z = \sigma_A(w, z) \tag{6.24a}$$

The expression $\sigma_A(z, w)$, which is shorthand for the covariance between the breeding values for the trait and relative fitness $\sigma(A_z, A_w)$, is often called the **Robertson covariance**. The term "secondary theorem" is reasonable for this version, as it directly follows from Fisher's fundamental theorem using a simple regression argument (Falconer 1985). The expected change in a trait is given by its change in mean breeding value ΔA_z, which in turn is the change in the breeding value of relative fitness ΔA_w times the regression of the trait breeding value A_z on A_w, or

$$R_z = \Delta A_z = \beta_{A_z, A_w} \, \Delta A_w \tag{6.24b}$$

From standard univariate regression theory (LW Chapter 3), the slope of the regression of A_z on A_w is just

$$\beta_{A_z,A_w} = \frac{\sigma(A_z, A_w)}{\sigma^2(A_w)} \tag{6.24c}$$

Likewise, under the fundamental theorem, $\Delta A_w = \sigma^2(A_w)$. Substituting these two results into Equation 6.24b recovers Robertson's 1968 version of his theorem,

$$R_z = \left(\frac{\sigma(A_z, A_w)}{\sigma^2(A_w)} \right) \sigma^2(A_w) = \sigma(A_z, A_w)$$

1966 Version: $R_z = \sigma(A_z, w)$

Formally, Robertson coined the term secondary theorem for his 1968 version. However, an (attempted) careful reading of Robertson's rather opaque earlier 1966 paper suggests his original result was a natural extension of the Robertson-Price identity, with

$$R_z = \sigma(w, A_z) \tag{6.25a}$$

being the within-generation change in the breeding value of the trait caused by selection, and hence the expected response in the next generation when average parental breeding value predicts offspring mean. Equation 6.25a is widely used in evolutionary quantitative genetics (e.g., Lande 1976; Frank 1997), and is closely connected with the breeder's equation. Taking the regression of breeding value on phenotype as $A_z = h^2(z - \mu) + e_A$ and substituting into Equation 6.25a gives

$$R_z = \sigma(w, h^2[z - \mu] + e_A) = h^2\sigma(w, z) + \sigma(w, e_A) = h^2 S + \sigma(w, e_A) \tag{6.25b}$$

and we recover the breeder's equation when $\sigma(w, e_A) = 0$. Biologically, this latter contribution means that the relative fitness of an individual is uncorrelated with the residual in estimating their breeding value from phenotype. This term can be nonzero when an environmental factor influences both fitness and trait value, in which case the breeder's equation can fail, while Robertson's theorem may still hold (Chapter 20). Writing $w = A_w + e_w$,

$$\sigma(w, A_z) = \sigma(A_w + e_w, A_z) = \sigma(A_w, A_z) + \sigma(e_w, A_z) \tag{6.25c}$$

showing that while the 1966 version is more general, the two are equal when $\sigma(e_w, A_z) = 0$.

Accuracy of the Secondary Theorem

Formal population-genetic analysis of the secondary theorem assumes the 1968 version, and is very closely related to work on the accuracy of the classical interpretation of Fisher's theorem. Recall Equation 5.22, which gave the contribution to selection response from the kth locus (in the absence of epistasis) as

$$R_k = 2\sum_j \alpha_j s_j p_j + \sum_{i,j} \delta_{ij} p_i s_i p_j s_j$$

where all terms relate to the values for locus k. The first sum is the expected product of the average effect α_j of an allele on character value times the average excess s_i of that allele on relative fitness. Recall (LW Equation 3.8) that $\sigma(x, y) = E[x \cdot y] - E[x] \cdot E[y]$. Since (by definition), $E[\alpha_j] = E[s_i] = 0$, the first sum is the covariance between the average effect of the character with the average excess on fitness, in other words, the additive genetic covariance

between relative fitness and the focal trait. Summing over all loci, we can express Equation 5.22 as

$$R = \sigma_A(z, w) + B \tag{6.26}$$

If the character has no dominance (all $\delta_{ij} = 0$), the correction term B vanishes, recovering Robertson's original suggestion. Even if dominance is present, in the absence of epistatic variance, the error in the secondary theorem becomes increasingly small as the number of loci increases (Example 5.12).

The most general statement on the validity of Robertson's 1968 version is due to Nagylaki (1992b, 1993) and assumes weak selection on the underlying loci and random mating, but allows for arbitrary epistasis and linkage disequilibrium. Similar to his weak selection analysis of Fisher's theorem discussed above, Nagylaki shows that after a sufficient time the change in mean trait value is given by

$$R = \sigma_A(z, w) + O(s^2) \tag{6.27}$$

As with the fundamental theorem, when gametic frequencies approach their equilibrium values, terms of order s^2 can become significant and mean response can differ significantly from Robertson's prediction, but again, the bulk of evolutionary change likely occurs before we approach the equilibrium value too closely. Consequently, the amount of change during the final approach to the equilibrium is generally expected to be quite small, so that (as with the fundamental theorem) under weak selection on the underlying loci, Robertson's theorem holds for the bulk of evolutionary change.

Connecting Robertson's Results with those of Price, Fisher, and Lush

As we have shown, Robertson's 1968 version directly follows from the classical version of Fisher's theorem, while his 1966 version easily recovers the breeder's equation. Recall that we also showed that the breeder's equation recovers the classical version of Fisher's theorem (Equation 6.17b). All that remains is to consider how Robertson's results fit with Price's theorem. Following the derviation leading to Equation 6.20b, let z in Equation 6.7c denote the breeding value of the trait of interest. The between-generation change in the mean breeding value for our focal trait is

$$\begin{aligned} \Delta\overline{A}_z &= \sigma(w, A_z + \overline{\delta}) + E(\overline{\delta}) \\ &= \sigma(w, A_z) + \sigma(w, \overline{\delta}) + E(\overline{\delta}) \end{aligned} \tag{6.28a}$$

showing that Robertson's (1966) version naturally follows from Price's theorem. As with our analysis of Fisher's theorem within the Price equation framework, if parental breeding values exactly predict mean offspring breeding values, then $\overline{\delta} = 0$ and Robertson's equation holds exactly. When parental breeding values are good predictors of mean offspring breeding value, so that $\overline{\delta}$ is not zero but still small, then Roberston's (1966) theorem is a good approximation for response. Writing $w = A_w + e_w$,

$$\begin{aligned} \Delta\overline{A}_z &= \sigma(A_w + e_w, A_z) + \sigma(w, \overline{\delta}) + E(\overline{\delta}) \\ &= \sigma(A_w, A_z) + \sigma(e_w, A_z) + \sigma(w, \overline{\delta}) + E(\overline{\delta}), \end{aligned} \tag{6.28b}$$

which recovers Robertson's 1968 version, provided that the last three terms are zero (or very small relative to the first term).

THE BREEDER'S EQUATION VIEWED FROM THE PRICE EQUATION

The breeder's equation is the backbone for much of the theory of selection response in quantitative genetics. Almost all of the theory of breeding is framed around it, as is much

of evolutionary quantitative genetics, although Robertson's secondary theorem has been gaining some recent traction in studies in natural populations (reviewed in Chapter 20). As the above results show, the breeder's equation is an approximation, with a linear parent-offspring regression a necessary, but not sufficient, condition (Example 6.1). Viewed in a population-genetic framework, Example 5.12 showed that this approximation is generally fairly good under weak selection on each of the loci underlying a trait.

Example 6.6. Consider a single locus that completely determines the value of a trait, where genotypes **QQ** and **Qq** have phenotypic value one, while **qq** has value zero (there is no environmental variance), and **qq** individuals have a survival rate twice as high as **QQ/Qq** individuals. Let's contrast the exact response from single-locus models (Chapter 5) with that predicted by the breeder's equation. The presence of a single major gene with dominance is expected to lead to poor performance of the latter. Starting with the exact population genetics model, letting p be the frequency of allele **q**,

Genotype	**QQ**	**Qq**	**qq**
Trait value	1	1	0
Frequency (before selection)	$(1-p)^2$	$2p(1-p)$	p^2
Fitness	1	1	2
Frequency (after selection)	$(1-p)^2/\overline{W}$	$2p(1-p)/\overline{W}$	$p^2(2/\overline{W})$

The resulting trait mean before selection is $\mu(p) = 1-p^2$, with mean fitness $\overline{W}(p) = 1+p^2$. The new frequency p' of allele **q** after selection is half the frequency of **Qq** after selection plus the (after-selection) frequency of **qq**,

$$p' = \frac{(1/2)2p(1-p)+2p^2}{1+p^2} = \frac{p(1+p)}{1+p^2}, \tag{6.29a}$$

for an allele frequency change of

$$\Delta p = p' - p = \frac{p(1+p)}{1+p^2} - p\,\frac{1+p^2}{1+p^2} = \frac{p^2(1-p)}{1+p^2},$$

which translates into a change in mean phenotype of

$$R = \mu(p') - \mu(p) = p^2 - (p')^2 = -2p^3\frac{(1-p(1+p^2)/2}{(1+p^2)^2}. \tag{6.29b}$$

This exact single-generation response in the trait mean is plotted below.

Now consider the response predicted from the breeder's equation. Using the standard trait-value parameterization of $2a : a(1+k) : 0$, we have $a = 1/2$ and $k = 1$, giving (LW Chapter 4) the additive variance of the trait as

$$\sigma_A^2 = 2p(1-p)a^2[1+k(2p-1)]^2 = 2p(1-p)(1/4)[2p]^2 = 2p^3(1-p)$$

$$\sigma_D^2 = (2p(1-p)ak)^2 = p^2(1-p)^2$$

Since σ_E^2 is assumed to be zero, the heritablity becomes

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2+\sigma_D^2} = \frac{2p^3(1-p)}{2p^3(1-p)+p^2(1-p)^2} = \frac{2p}{1+p}. \tag{6.29c}$$

Following selection, the fitness-weighted mean is

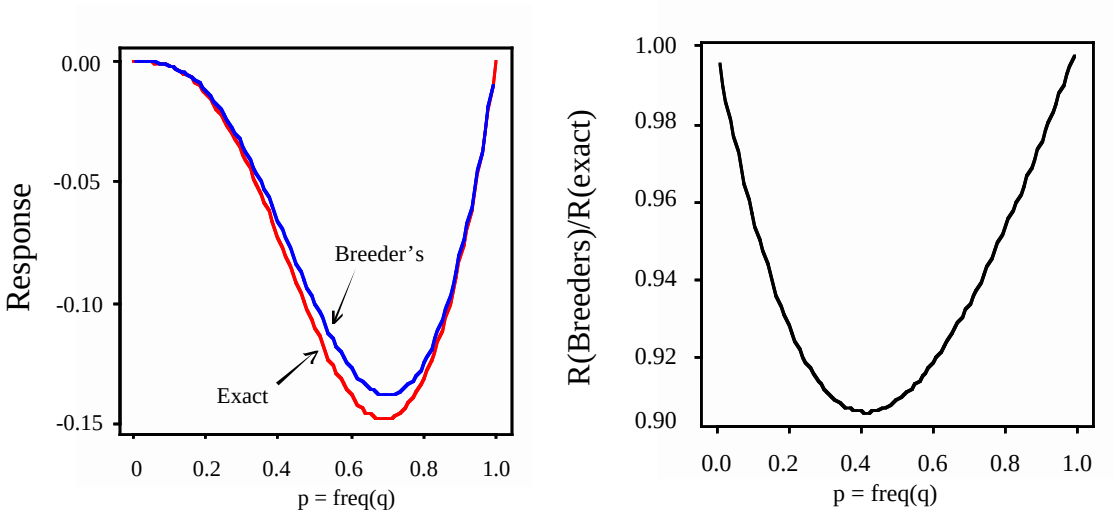
$$\mu' = 1 \cdot \frac{1}{\overline{W}}\left(1-p^2\right) + 0 \cdot \frac{2}{\overline{W}}\,p^2 = \frac{1-p^2}{\overline{W}} = \frac{1-p^2}{1+p^2}, \tag{6.29d}$$

giving the selectional differential as

$$S = \mu' - \mu = \frac{1-p^2}{1+p^2} - \left(1-p^2\right) = -p^2\,\frac{1-p^2}{1+p^2}. \tag{6.29e}$$

Equations 6.29c and e give the predicted response from the breeder's equation as

$$R = h^2 S = -2p^3\,\frac{1-p^2}{(1+p^2)(1+p)}. \tag{6.29f}$$



Plotting the exact response (Equation 6.29b) and that predicted by the breeder's equation (Equation 6.29f), we see that the breeder's equation generally does well, but slightly underestimates the response, predicting (in the worst case) only about 90% of the actual response when $p \simeq 0.4$. What accounts for this discrepancy? For this simple one-locus model, two factors. First, the parent-offspring regression is *not* linear due to dominance. Second, the change in allele frequency in the selected parents results in changes the parent-offspring covariance, and hence the parent-offspring regression slope.

Beyond the Breeder's Equation: Heywood's Decomposition

When the parent-offspring regression is linear, Equation 6.12 shows that we recover the breeder's equation plus two correction terms. What can be said if we make no assumptions about the functional form of this regression? The most general solution is **Heywood's decomposition** (2005), based on the Price equation. This yields components that have biologically meaningful interpretations, rather than simply being mathematically convenient correction terms. A key point from this decomposition was foreshadowed in Example 6.6 — it is the parent-offspring regression *after*, as opposed to before, selection that is more appropriate for predicting response.

To express response in terms of the Price equation, as above let z_i denote the value of the ith parent/midparent whose offspring have mean value $\overline{z}_i$ and parent-offspring deviation $\overline{\delta}_i = \overline{z}_i - z_i$. Equation 6.7c gives the general expression for response as

$$R = \sigma(w_i, \overline{z}_i) + E(\,\overline{\delta}_i\,) \tag{6.30a}$$

Heywood's key insight is that Equation 6.30a can be decomposed into contributions from two biologically meaningful linear regressions and their corresponding residuals. The first is the regression of offspring value $\overline{z}_i$ on parental value z_i,

$$\overline{z}_i = \mu + \beta_{\overline{z},z}\, z_i + e_{\,\overline{z},i}, \quad \text{where} \quad \beta_{\overline{z},z} = \frac{\sigma(\overline{z}_i, z_i)}{\sigma_z^2} \tag{6.30b}$$

The second is the regression of relative fitness w of a parent on its phenotypic value z,

$$w_i = 1 + \beta_{w,z}\, z_i + e_{w,i}, \quad \text{where} \quad \beta_{w,z} = \frac{\sigma(w_i, z_i)}{\sigma_z^2} = \frac{S}{\sigma_z^2} \tag{6.30c}$$

It is important to stress that we are *not* assuming that the true relationship between either the mean offspring value $\overline{z}_i$ or paternal fitness w_i and paternal phenotype value z_i is linear, rather simply considering the best linear regression for these relationships and the resulting residuals.

The Partial Covariance and the Spurious Response to Selection

An important idea in statistic is the residual dependency between two variables (say x and y) once the effects of a third (z) is removed. One dependency measure is the **conditional covariance**, $\sigma(x, y \mid z)$, and is defined as one might expect, namely the covariance conditioned on a partial value of z,

$$\sigma(x, y \mid z) = E\left[\left(x - E[x|z]\right)\left(y - E[y|z]\right) \mid z\right] \tag{6.31a}$$

While this might seem the obvious measure to use, the one that appears in Heywood's decomposition is the **partial covariance,**

$$\sigma(x, y \cdot z) = \sigma(e_x, e_y) \tag{6.31b}$$

where e_x and e_y are the residuals for the linear regression of x on z and of y on z. If the three variables are multivariate normal, the conditional and partial covariances are equal, but generally they differ (Lawrance 1976; Baba et al. 2006). The partial covariance can be expressed in terms of the various pairwise covariances between the three variables as follows. The linear regression of x on z can be written as

$$x = \mu_x + \beta_{x,z} z + e_x = \mu_x + \left(\frac{\sigma(x,z)}{\sigma_z^2}\right) z + e_x,$$

implying that

$$e_x = x - (\mu_x + \beta_{x,z}\, z) = x - \mu_x - \left(\frac{\sigma(x,z)}{\sigma_z^2}\right) z \tag{6.31c}$$

with a similar expression for e_y. Substituting these into Equation 6.31b and ignoring the mean terms (which, as constants, do not factor in the covariance) gives

$$\sigma(x, y \cdot z) = \sigma(e_x, e_y) = \sigma\left(x - \left(\frac{\sigma(x,z)}{\sigma_z^2}\right) z, y - \left(\frac{\sigma(y,z)}{\sigma_z^2}\right) z\right)$$

Expanding the covariance yeilds

$$\begin{aligned}\sigma(x, y \cdot z) &= \sigma(x, y) - \left(\frac{\sigma(x, z)}{\sigma_z^2}\right)\sigma(y, z) - \left(\frac{\sigma(y, z)}{\sigma_z^2}\right)\sigma(x, z) + \left(\frac{\sigma(y, z)}{\sigma_z^2}\right)\left(\frac{\sigma(x, z)}{\sigma_z^2}\right)\sigma_z^2 \\ &= \sigma(x, y) - \left(\frac{\sigma(x, z)}{\sigma_z^2}\right)\sigma(y, z) \qquad (6.31\text{d})\end{aligned}$$

The last term could also be expressed as either $\beta_{x,z}\sigma(y, z)$ or $\beta_{y,z}\sigma(x, z)$. A key component in Heywood's decomposition is the partial covariance $\sigma(\overline{z}, w \cdot z)$ between offspring mean $\overline{z}_i$ and parental fitness w_i given the trait value z_i of the parent. From Equation 6.31d,

$$\sigma(\overline{z}, w \cdot z) = \sigma(\overline{z}, w) - \frac{\sigma(\overline{z}, z)}{\sigma_z^2}\,\sigma(w, z) = \sigma(\overline{z}, w) - \beta_{\overline{z},z}\,\sigma(w, z) \qquad (6.31\text{e})$$

Rearranging this last expression gives

$$\begin{aligned}\sigma(\overline{z}, w) &= \sigma(\overline{z}, w \cdot z) + \beta_{\overline{z},z}\,\sigma(w, z) \\ &= \sigma(\overline{z}, w \cdot z) + \beta_{\overline{z},z}\,S, \qquad (6.31\text{f})\end{aligned}$$

with substitution into Equation 6.30a yielding

$$R = \beta_{\overline{z},z}\,S + \sigma(w, \overline{z} \cdot z) + E(\overline{\delta}) \qquad (6.32)$$

Thus, even when the parent-offspring regression is non-linear, we can express response as a linear response (breeder's equation) term $\beta_{\overline{z},z}S$ plus a correction $\sigma(\overline{z}, w \cdot z) = \sigma(e_{\overline{z}}, e_w)$ accounting for (among other things) non-linearity and a second correction $E(\overline{\delta})$ for transmission.

Comparing Equations 6.12 and 6.32, we see that $\sigma(w_i, e_i) = \sigma(w, \overline{z} \cdot z)$, so that the term expressed as a simple covariance in Equation 6.12 can also be defined by a partial covariance given z. Since $\sigma(w, \overline{z} \cdot z)$ removes the (linear) effect of parental value z_i on both its offspring mean $\overline{z}_i$ and its own relative fitness w_i, any residual association between w_i and $\overline{z}_i$ is uncorrelated with z_i and hence uncorrelated with selection on the parent. Thus, the response from this component need not be adaptive, leading Heywood to denote this as the **spurious response to selection**.

There are two ways to generate a nonzero $\sigma(w, \overline{z} \cdot z)$. First, both regressions (w_i on z_i and $\overline{z}_i$ on z_i) may be nonlinear and, as a result, their residuals may be correlated, generating a spurious response (e.g., Figure 6.3). Second, even if one (or both) of these regressions are linear, if both w_i and $\overline{z}_i$ are correlated through an unmeasured third variable (such as an environmental effect), their residuals after being regressed on z_i can still be correlated, again generating a potentially spurious response (Chapter 20 examines this in more detail).

Parent-Offspring Regressions Before and After Selection

A subtle, but important, point to stress about Equation 6.32 is that $\beta_{\overline{z},z}$ is the slope of the regression of offspring mean *following* selection. As first suggested in Example 6.6, this slope may be different from the regression based upon unselected parents. In particular, when considering the mean offspring value given a *single* parent, selection can change the mean of their offspring by changing the distribution of genotypes for their mates. This is *not* the case when we take the unit to be the midparent, as the distribution of the offspring from the midparent is independent of the values of *other* midparents.The same is true for parents that asexually reproduce or that self. For these cases, Equation 6.32 provides the extension of the breeder's equation to general settings. A potential exception, examined in Chapter 22, is when the trait value is a function of both the individual and the members of its group.

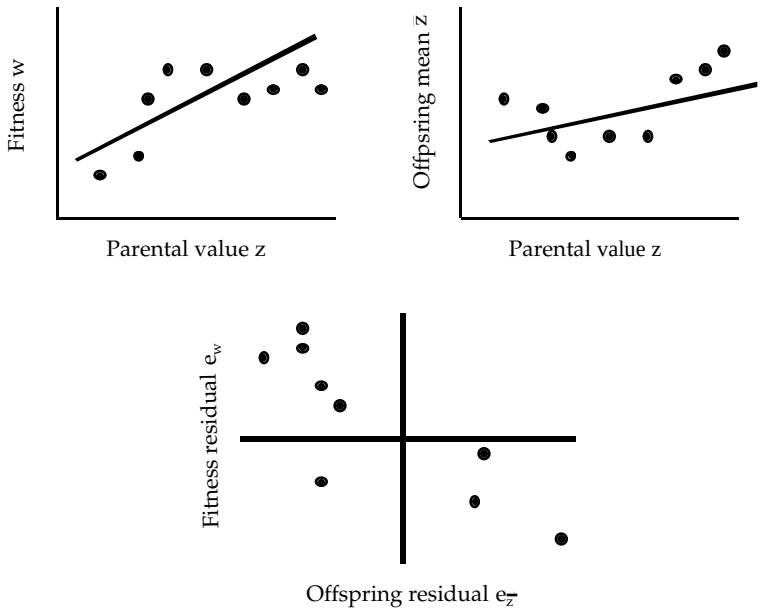


Figure 6.3. The partial covariance $\sigma(\overline{z}, w \cdot z)$ is the covariance $\sigma(e_{\overline{z}}, e_w)$ between the residuals of the regression of w_i on z_i and the regression of $\overline{z}_i$ on z_i (parent-offspring regression). In this example, both of these relationships are nonlinear, resulting in the distribution of residuals for the best fitting linear regression being non-randomly distributed. A plot of these residuals against each other shows a strong negative covariance and hence while $\sigma(w, z)$, $\sigma(\overline{z}, z)$, and $\sigma(w, \overline{z})$, are all positive, $\sigma(w, \overline{z} \cdot z)$ is negative.

Thus, when considering single parents, Equation 6.32 needs to be modified to account for the possibility that the mean offspring from this parent is influenced by the distribution of other parents after selection. To do so, let $\overline{z}_i$ and $\overline{z}_i^*$ denote the offspring mean from parent i when mated to parents from selected and unselected populations, respectively. Given that Equation 6.32 is expressed in terms of $\overline{z}_i$, following Heywood we wish to translate this equation into an expression involving $\overline{z}_i^*$ (the offspring means of the parents in the absence of selection). Finally, define $\delta_i^* = \overline{z}_i - \overline{z}_i^*$ as the difference in the offspring means for parent i when crossed to selected versus unselected parents. Since $\overline{z}_i = \overline{z}_i^* + \delta_i^*$,

$$\sigma(\overline{z}_i, z_i) = \sigma(\overline{z}_i^* + \delta_i^*, z_i) = \sigma(\overline{z}_i^*, z_i) + \sigma(\delta_i^*, z_i) \tag{6.33}$$

Thus

$$\beta_{\overline{z},z} = \frac{\sigma(\overline{z}, z)}{\sigma_z^2} = \frac{\sigma(\overline{z}^*, z) + \sigma(\delta^*, z)}{\sigma_z^2} = \beta_{\overline{z}^*,z} + \frac{\sigma(\delta^*, z)}{\sigma_z^2} \tag{6.34}$$

Equation 6.34 relates the parent-offspring regressions using unselected ($\beta_{\overline{z}^*,z}$) versus selected ($\beta_{\overline{z},z}$) parents. The same logic leading to Equation 6.31f can be used to obtain two useful identities:

$$\sigma(w, \overline{z}^*) = \sigma(w, \overline{z}^* \cdot z) + \beta_{\overline{z}^*,z}\sigma(w, z) \tag{6.35a}$$

and

$$\sigma(w, \overline{z} \cdot z) = \sigma(w, \overline{z}^* \cdot z) + \sigma(w, \delta^*) - \frac{\sigma(\delta^*, z)\,\sigma(w, z)}{\sigma_z^2} \tag{6.35b}$$

Finally, it will prove useful to decompose $\overline{\delta}_i = \overline{z}_i - z_i$. While $\overline{\delta}_i$ is often considered the change from parent to offspring in the absence of selection, this is not strictly correct. Formally, it has two components,

$$\overline{\delta}_i = \overline{z}_i - z_i = (\overline{z}_i^* + \delta_i^*) - z = \delta_i^* + (\overline{z}_i^* - z_i) \tag{6.36}$$

Hence

$$E(\,\overline{\delta}_i\,) = E(\,\delta_i^*\,) + E(\overline{z}_i^* - z_i) \tag{6.37}$$

where $E(\overline{z}_i^* - z_i)$, which Heywood calls the **constitutive transmission bias**, is the change from parent to offspring using *unselected* parents. Inbreeding, decay of linkage disequilibrium, and randomization of correlated environmental effects are all examples of situations with the potential for a non-zero constitutive transmission bias. The second component, $E(\,\delta_i^*\,)$, Heywood's **general induced transmission bias**, is the difference between the offspring means when the parent is crossed to the selected population versus the offspring mean when crossed to the unselected population.

Heywood's Decomposition of Response

We now have all of the components in place for Heywood's decomposition of the selection response. Expressed in terms of a regression using parents before selection ($\beta_{\overline{z}^*,z}$), substituting Equations 6.34 – 6.37 into Equation 6.32 yields

$$R = \beta_{\overline{z}^*,z}\, S + \sigma(w, \overline{z}^* \cdot z) + \sigma(w, \delta^*) + E(\,\delta^*\,) + E(\,\overline{z}^* - z) \tag{6.38}$$

Conversely, expressed in terms of the regression after selection ($\beta_{\overline{z},z}$), the response is

$$R = \beta_{\overline{z},z} S + \sigma(w, \overline{z}^* \cdot z) + \sigma(w, \delta^* \cdot z) + E(\,\delta^*\,) + E(\,\overline{z}^* - z\,). \tag{6.39}$$

This is obtained from Equation 6.38, using Equation 6.34 and the identity

$$\sigma(w, \delta^*) - \frac{\sigma(\delta^*, z)\,\sigma(w, z)}{\sigma_z^2} = \sigma(w, \delta^* \cdot z) \tag{6.40}$$

Equation 6.40 follows using the same logic leading to Equation 6.31f. Besides the different regression coefficients ($\beta_{\overline{z}^*,z}$ vs $\beta_{\overline{z},z}$), Equations 6.38 and 6.39 differ in that $\sigma(w, \delta^*)$ in Equation 6.38 is replaced by $\sigma(w, \delta^* \cdot z)$ when we use the regression on selected parents.

Heywood notes that the last four terms in Equations 6.38 and 6.39 arise because of imperfect transmission of parental to offspring phenotype. If transmission is perfect, then $z_i = \overline{z}_i^* = \overline{z}_i$ and $\delta_i^* = 0$, and these last four terms are all zero. Although all four terms are due to imperfect transmission, they have different causes and can be thought of as logically distinct processes. Heywood denotes the different contributions to response in Equations 6.38 and 6.39 as follows:

Term	Heywood's Interpretation
βS	Linear response to selection
$\sigma(w, \overline{z}^* \cdot z)$	Spurious response to selection
$\sigma(w, \delta^*), \sigma(w, \delta^* \cdot z)$	Special induced transmission bias
$E(\delta^*)$	General induced transmission bias
$E(\,\overline{z}^* - z\,)$	Constitutive transmission bias

The linear response to selection is the breeder's equation analog. In the decomposition given by Equation 6.38 this term is exactly the breeder's equation (assuming that $\beta_{\overline{z}^*,z} = h^2$) as

the regression coefficient is based on unselected parents. In the version given by Equation 6.32, β is the parent-offspring regression following selection. This regression may change as the fitness scheme changes, while the before-selection regression remains unchanged under any fitness scheme.

Example 6.7. Heywood offered the following example to illustrate the roles played by the various components in his decomposition. Consider the situation where trait value z is determined entirely by a single diallelic locus and fitness is a simple linear function of z, $W = 1 + t\,z$. If μ denotes the mean trait value before selection, then mean fitness is just $\overline{W} = 1 + t\,\mu$, giving the relative fitness as

$$w = W/\overline{W} = \frac{1 + t\,z}{1 + t\,\mu}$$

The resulting selection differental and mean after selection (but before reproduction) are

$$S = \sigma(w, z) = \frac{t}{1 + t\mu}\,\sigma(z, z) = \sigma_z^2\tau, \qquad \mu_s = \mu + S = \frac{\mu + t\,(\sigma_z^2 + \mu^2)}{1 + t\mu}$$

where $\tau = t/(1 + t\mu) = S/\sigma_z^2$. The breeder's equation predicts the response as

$$R_B = h^2 S = \sigma_A^2\tau$$

However, the actual response (after a bit of algebra following the same logic as in Example 6.6) is found to be

$$R_A = \sigma_A^2\tau - (1/2)\tau^2\sigma_A^2\sigma_D = h^2 S\left(1 - S\,\frac{\sigma_D}{2\sigma_z^2}\right)$$

where σ_D^2 is the dominance variance.

Let's examine Heywood's decomposition in this case. First, with a single locus under random mating, the offspring mean from a collection of unselected parents is the same as the parental mean, giving $E(\overline{z}^* - z) = 0$. However, the difference in the mean between the offspring of a parent crossed to a collection of unselected parents and a parent crossed to the selected parents is given by

$$E(\,\delta^*\,) = E(\overline{z} - z^*) = \frac{\sigma_A^2}{2}\,\tau = \frac{h^2 S}{2},$$

which is obtained by computing the values of $\overline{z}_i$ for this fitness scheme. Likewise, the parent-offspring covariances can be found to be

$$\sigma(\overline{z}^*, z) = \frac{\sigma_A^2}{2}, \qquad \sigma(\overline{z}_i, z_i) = \frac{\sigma_A^2}{2}\,(1 - \sigma_D\tau)$$

As expected, the (unselected) parent-offspring covariance $\sigma(\overline{z}^*, z)$ is just $\sigma_A^2/2$. However, when our parent is crossed to a parent from the selected population, the resulting parent-offspring covariance $\sigma(\overline{z}_i, z_i)$ differs from this when $\sigma_D^2 \neq 0$ (i.e., dominance is present).

Finally, since $w_i = c + z_i\tau$ for a constant c independent of z, the covariances between offspring means and parental fitness are

$$\sigma(w_i, \overline{z}_i) = \tau\,\sigma(\overline{z}_i, z_i), \qquad \sigma(w_i, \overline{z}_i^*) = \tau\,\sigma(\overline{z}_i^*, z_i)$$

Now let's compute the remaining components in Heywood's decomposition. Consider first the linear response to selection. Using the regression before selection recovers the breeder's equation for this component,

$$\beta_{\overline{z}^*,z} S = \frac{\sigma(\overline{z}^*, z)}{\sigma_z^2}\, S = \frac{\sigma_A^2/2}{\sigma_z^2}\, S = \frac{h^2 S}{2}$$

In contrast, for the regression when the parent mates with an individual from the selected population, the linear response is

$$\beta_{\overline{z},z} S = \frac{\sigma(\overline{z}, z)}{\sigma_z^2} S = \frac{\sigma_A^2}{2\sigma_z^2}\, S\,(1 - \sigma_D\tau) = \frac{h^2 S}{2}\left(1 - S\,\frac{\sigma_D}{\sigma_z^2}\right)$$

The spurious response to selection, $\sigma(w, \overline{z}^* \cdot z)$, is obtained by recalling $\sigma(w, z) = S = \tau\,\sigma_z^2$ and using Equation 6.35a to give

$$\begin{aligned}\sigma(w, \overline{z}^* \cdot z) &= \sigma(w, \overline{z}^*) - \beta_{\overline{z}^*,z}\sigma(w, z) \\ &= \tau\,\sigma(z, \overline{z}^*) - \frac{\sigma(z, \overline{z}^*)}{\sigma_z^2}\,\tau\,\sigma_z^2 = 0\end{aligned}$$

Lastly, consider the special induced transmission bias. When we use the regression on unselected parents (Equation 6.38), the induced transmission bias is given by $\sigma(w, \delta^*) = \sigma(w, \overline{z} - \overline{z}^*)$,

$$\begin{aligned}\sigma(w, \delta^*) &= \sigma(w, \overline{z}) - \sigma(w, \overline{z}^*) = \tau\,[\,\sigma(\overline{z}, z) - \sigma(\overline{z}^*, z)\,] \\ &= -\frac{\tau^2\sigma_A^2\sigma_D}{2} = -\,h^2\,S^2\,\frac{\sigma_D}{2\sigma_z^2}\end{aligned}$$

When we use the decomposition based on the regression of selected parents (Equation 6.39), the special induced bias is given by $\sigma(w, \delta^* \cdot z)$, which from Equation 6.40 gives

$$\sigma(w, \delta^* \cdot z) = \sigma(w, \delta^*) - \frac{\sigma(\delta^*, z)\,\sigma(w, z)}{\sigma_z^2}$$

Noting that

$$\sigma(\delta^*, z) = \sigma(\overline{z} - \overline{z}^*, z) = \sigma(\overline{z}, z) - \sigma(\overline{z}^*, z) = -\frac{\sigma_A^2\sigma_D\tau}{2}$$

we have

$$\sigma(w, \delta^* \cdot z) = -\frac{\tau^2\sigma_A^2\sigma_D}{2} - \left(-\frac{\sigma_A^2\sigma_D\tau}{2}\right)\left(\frac{\tau\sigma_z^2}{\sigma_z^2}\right) = 0$$

Summarizing the above results, the total response can be decomposed two different ways, depending on whether we use an unselected parent-offspring regression (Equation 6.38) or the regression using selected parents (Equation 6.39) as our frame of reference,

Component	Pre-selection regression	Post-selection regression
Linear Response	$(h^2/2)\,S$	$(h^2/2)(1 - S\sigma_D/\sigma_z^2)\,S$
Spurious response	0	0
Special bias	$-h^2S^2\,\sigma_D/(2\sigma_z^2)$	0
General bias	$h^2S/2$	$h^2S/2$
Constitutive bias	0	0

A couple of points are immediately apparent. First, slopes of the linear regressions are different (for $\sigma^2_D > 0$) when we cross our parent to unselected vs. selected parents. Thus, if we used the parent-offspring regression estimated using unselected parents as the basis for our predicted response, we would overestimate response when dominance variance is present.

Second, there is no constitutive transmission bias, as under the mating scheme assumed (random mating), in the absence of fitness variation, the mean of the offspring equals the mean of the parents and hence $E(\overline{z}^* - z) = 0$. However, note that the general induced transmission bias is the same under the two types of regressions. This bias is $E(\,\delta'\,) = E(\overline{z} - \overline{z}^*)$, the difference in the expected value of offspring when the parent is crossed to selected parents as opposed to unselected parents. This makes perfect sense, as the mean breeding values in the selected parents is $h^2 S$, thus their additional contribution to the offspring is half their breeding value, $h^2 S/2$. (With unselected parents, the mean breeding value is zero and hence no additional contribution.)

The special induced transmission bias when using a regression before selection, $\sigma(w, \delta^*)$, is nonzero, with this term accounting for the reduction from the breeder's equation by not considering that the parental regression changes when crossed to a selected parent. When considering the regression after selection, the special bias is given by $\sigma(w, \delta^* \cdot z)$, which in this case is zero as we have accounted for the change in regressions.

Finally, in both cases the spurious response is zero. Recall that a non-zero spurious response can arise in two settings. First, both the parent-offspring ($\overline{z}_i$ on z_i) and parent-fitness (w_i on z_i) regressions are non-linear, in which case the residuals from both regressions may be correlated. Second, even if both regressions are linear, residuals can still be correlated through some additional factor, such as an environmental influence on both $\overline{z}_i$ and w_i, that we have not accounted for. The latter does not apply for this simple problem, and even though the parent-offspring regression is nonlinear (with dominance), the parent-fitness regression $(1 + tz)$ is linear in this case. Under more general settings, we would expect this fitness regression to be nonlinear, potentially generating a spurious response.

Example 6.8. Example 6.7 was based on single parent-offspring regressions, so that the mean of the offspring from a given individual changes when we (hypothetically) cross that individual to a random member from the selected vs. unselected populations. Now suppose we take the ancestral unit as the midparent. In this case, the mean value of the offspring, given the midparent, is completely unaffected by the distribution of all other genotypes in the selected population. This occurs because the offspring mean only depends upon the genotypes of its two parents, and is normally unaffected by the distribution of the genotypes of any other individuals (but see Chapter 22). In this case, $\overline{z}_i = \overline{z}_i^*$ (the mean of offspring from this pair in the context of a selected population of other midparents is the same as in the context of an unselected population of other midparents), and response is given by Equation 6.32,

$$R = \beta_{\overline{z},z}\, S + \sigma(w, \overline{z} \cdot z) + E(\,\overline{\delta}\,)$$

Further, since $\delta_i^* = \overline{z}_i - \overline{z}_i^* = 0$,

$$E(\,\overline{\delta}\,) = E(\delta^*) + E(\overline{z}^* - z) = 0 + 0 = 0$$

Likewise, the slope of the midparent-offspring regression is just h^2, giving

$$R = h^2\, S + \sigma(w, \overline{z} \cdot z_{mp})$$

In this case, unlike Example 6.6, the regression of fitness on midparent value is *nonlinear*, as the fitness of the midparent pair (z_p, z_m) is the product of their fitnesses,

$$W(z_p) \cdot W(z_m) = (1 - tz_p)(1 - tz_m) = 1 - 2tz_{mp} + t^2 z_m z_p$$

Thus, both the midparent-offspring and midparent-fitness regressions are nonlinear, and there is a correlation between the residuals of these two regressions (see Heywood 2005 for details),

$$\sigma(w, \overline{z} \cdot z) = -h^2 S^2 \frac{\sigma_D}{2\,\sigma_z^2}$$

This generates a spurious response, giving total response from the framework of midparent-offspring regression as

$$R = h^2\, S - h^2 S^2 \frac{\sigma_D}{2\,\sigma_z^2}$$

Which, as expected, is the same total response as when we considered response from two (pre- and post-selection) single parent-offspring regression frameworks.

These last two examples highlight a second utility of Heywood's decomposition, looking at the components of response under different frames of reference. When considered from a single individual versus midparent framework for the ancestors, we obtain slightly different decompositions — the sum (total response) is the same, but the distribution of components differs. Using the individual as the framework (and hence the fitness of individuals), there is no spurious component of selection response, as selection is based on a linear function of individual phenotype. In contrast, when considered from the midparental perspective (and hence the fitness of the midparent), although the total response remains the same, there *is* a spurious component, because the linear individual-based fitness function becomes a nonlinear midparental fitness function. Considering the components of response from different ancestral frameworks can provide useful insight into an evolutionary problem, see Heywood for additional examples.

Literature Cited

Baba, K., R. Shibata, and M. Sibuya. 2005. Partial correlation and conditional correlation as measures of conditional independence. *Aust. N. Z. J. Stat.* 46: 657–664. [6]

Boyd, R., and P. Richerson. 2005. *The origin and evolution of cultures.* Oxford University Press, Oxford, UK. [6]

Campbell, D. R. 1997. Genetic and environmental variation in life-history traits of a monocarpic perennial: a decade-long field experiment. *Evolution* 51: 373–382. [6]

Charlesworth, B. 1984. The evolutionary genetics of life histories. *In* B. Shorrocks (ed.) *Evolutionary ecology*, pp. 117–133. Blackwell Scientific, Oxford, UK. [6]

Charlesworth, B. 1987. The heritability of fitness. *In* J. W. Bradbury and M. B. Andersson (eds.) *Sexual selection: testing the alternatives,* pp. 21-40. Wiley, New York, NY. [6]

Coltman, D. W., P. O'Donoghue, J. T. Hogg, and M. Festa-Bianchet. 2005. Selection and genetic (co)variance in bighorn sheep. *Evolution* 59: 1372–1283. [6]

Crnokrak, P., and D. A. Roff. 1995. Dominance variance: associations with selection and fitnesses. *Heredity* 75: 530–540. [6]

Crow, J. F., and T. Nagylaki. 1976. The rate of change of a character correlated with fitness. *Amer. Nat.* 110: 207–213. [6]

DeRose, M. A., and D. A. Roff. 1999. A comparison of inbreeding depression in life-history and morphological traits in animals. *Evolution* 53: 1288–1292. [6]

Edwards, A. W. F. 1990. Fisher, $\overline{W}$, and the fundamental theorem. *Theor. Popul. Biol.* 38: 276–284. [6]

Edwards, A. W. F. 1994. The fundamental theorem of natural selection. *Biol. Rev.* 69: 443–474. [6]

Ewens, W. J. 1969. A generalized fundamental theorem of natural selection. *Genetics* 63: 531–537. [6]

Ewens, W. J. 1976. Remarks on the evolutionary effect of natural selection. *Genetics* 83: 601–607. [6]

Ewens, W. J. 1989b. An interpretation and proof of the fundamental theorem of natural selection. *Theor. Popul. Biol.* 36: 167–180. [6]

Ewens, W. J. 1992. An optimizing principle of natural selection in evolutionary population genetics. *Theor. Popul. Biol.* 42: 333–346. [6]

Ewens, W. J. and G. Thompson. 1977. Properties of equilbria in multi-locus genetic systems. *Genetics* 87: 807–819. [6]

Ewens, W. J. 1994. The changing role of population genetics theory. *In* S. Levin (ed.), *Frontiers in mathematical biology,* pp. 186-197. Springer, New York, NY [6]

Falconer, D. S. 1985. A note on Fisher's 'average effect' and 'average excess'. *Genet. Res.* 46: 337–347. [6]

Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon, Oxford, UK. [6]

Fisher, R. A. 1941. Average excess and average effect of a gene substitution. *Ann. Eugen.* 11: 53–63. [6]

Fowler, K., C. Semple, N. H. Barton, and L. Partridge. 1997. Genetic variation for total fitness in *Drosophila melanogaster. Proc. R. Soc. Lond.* B 264: 191–199. [6]

Frank, S. A. 1995. George Price's contributions to evolutionary genetics. *J. Theor. Biol* 175: 373–388. [6]

Frank, S. A. 1997. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. *Evolution* 51: 1712–1729. [6]

Frank, S. A. 1998. *Foundations of social evolution.* Princeton University Press, Princeton, NJ. [6]

Frank, S. A. 2012. Natural selection. IV. The Price equation. *J. Evol. Biol.* 25: 1002–1019. [6]

Frank, S. A., and M. Slatkin. 1992. Fisher's fundamental theorem of natural selection. *Trends Ecol. Evol.* 7: 92–95. [6]

Gardner, A. 2008. The Price equation. *Current Bio.* 18: R198–R202. [6]

Gardner, M. P., K. Fowler, N. H. Barton, and L. Partridge. 2005. Genetic variation for total fitness in *Drosophila melanogaster*: complex yet replicable patterns. *Genetics* 169: 1553–1571. [6]

Gustafsson, L. 1986. Lifetime reproductive success and heritability: empirical support for Fisher's fundamental theorem. *Amer. Nat* 128: 761–764. [6]

Hansen, T. F., C. Pélabon, and D. Houle. 2011. Heritability is not evolvability. *Evol. Biol.* 38: 258–277. [6]

Harman, O. 2011. *The price of altruism: George Price and the search for the origins of kindness* . Norton, New York, NY. [6]

Helanterä, H., and T. Uller. 2010. The Price equation and extended inheritance. *Philos. Theor. Biol.* 2: e101. [6]

Heywood, J. S. 2005. An exact form of the breeder's equation for the evolution of a quantitative trait under natural selection. *Evolution* 59: 2287–2298. [6]

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. *Genetics* 130: 195-204/ [6]

Houle, D., B. Morikawa, and M. Lynch. 1996. Comparing mutational variabilities. *Genetics* 143: 1467–1483. [6]

Kempthorne, O. 1957. *An introduction to genetic statistics*. Iowa State University Press, Ames, IA. [6]

Kerr, B., and P. Godfrey-Smith. 2008. Generalization of the Price equation for evolutionary change. *Evol.* 63: 531–536. [6]

Kimura, M. 1965a. Attainment of quasi-linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* 52: 875–890. [6]

Kruuk, L. E. B., T. H. Clutton-Brock, J. Slate, J. M. Pemberton, S. Brotherstone, and F. E. Guniness. 2000. Heritability of fitness in a wild mammal population. *Proc. Natl. Acad. Sci. USA* 97: 698–703. [6]

Lawrance, A. J. 1976. On conditional and partial correlation. *Amer. Stat.* 30: 146–149. [6]

Lessard, S. 1997. Fisher's fundamental theorem natural selection revisited. *Theor. Popul. Biol.* 52: 119–136. [6]

Lessard, S., and A.-M. Castilloux. 1995. The fundamental theorem of natural selection in Ewens' sense (case of fertility selection). *Genetics* 141: 733–742. [6]

Li , C. C. 1967. Fundamental theorem of natural selection. *Nature* 214: 505–506. [6]

Lush, J. L. 1937. *Animal breeding plans.* Iowa State University Press, Ames. [6]

Mackay, T. F. C. 1985. A quantitative genetic analysis of fitness and its components in *Drosophila melanogaster. Genet. Res.* 47: 59–70. [6]

Maynard Smith, J., and G. R. Price. 1973. The logic of animal conflict. *Nature* 246: 15 – 18. [6]

McCleery, R. H., R. A. Pettifor, P. Armbruster, K. Meyer, B. C. Sheldon, and C. M. Perrins. 2004. Components of variance underlying fitness in a natural population of the great tit *Parus major*. *Amer. Nat.* 164: E62–E72. [6]

Merilä, J., and B. C. Sheldon. 1999. Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity* 83: 103–109. [6]

Merilä, J., and B. C. Sheldon. 2000. Lifetime reproductive success and heritability in nature. *Amer. Nat.* 155: 301–310. [6]

Messina, F. K. 1993. Heritability and 'evolvability' of fitness components in *Callosobruchus maculatus*. *Heredity* 71: 623–629. [6]

Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. *Heredity* 59: 181–197. [6]

Nagylaki, T. 1976. The evolution of one- and two-locus systems. *Genetics* 83: 583–600. [6]

Nagylaki, T. 1977a. *Selection in one– and two–locus systems.* Springer–Verlag, Berlin, Germany. [6]

Nagylaki, T. 1977b. The evolution of one– and two–locus systems. II. *Genetics* 85: 347–354. [6]

Nagylaki, T. 1991. Error bounds for the fundamental and secondary theorems of natural selection. *Proc. Natl. Acad. Sci. USA* 88: 2402–2406. [6]

Nagylaki, T. 1992b. Rate of evolution of a quantitative character. *Proc. Natl. Acad. Sci. USA* 89: 8121–8124. [6]

Nagylaki, T. 1993. The evolution of multilocus systems under weak selection. *Genetics* 134: 627–647. [6]

Okasha, S. 2006. *Evolution and levels of selection.* Oxford University Press, Oxford, UK. [6]

Plutynski, A. 2006. What was Fisher's fundamental theorem of natural selection and what was it for? *Stud. Hist. Phil. Biol & Biomed. Sci.* 37: 59–82. [6]

Price, G. R. 1970. Selection and covariance. *Nature* 227: 520–521. [6]

Price, G. R. 1972a. Extension of covariance selection mathematics. *Ann. Hum. Genet.* 35: 485–490. [6]

Price, G. R. 1972b. Fisher's 'fundamental theorem' made clear. *Ann. Hum. Genet.* 36: 129 – 140. [6]

Price, T. D., and D. Schluter. 1991. On the low heritability of life history traits. *Evolution* 45: 853–861. [6]

Rice, S. H. 2004. *Evolutionary theory: mathematical and conceptual foundations.* Sinauer Assocates, Inc. Sunderland, MA. [6]

Robertson, A. 1955. Selection in animals: synthesis. *Cold Spring Harbor Symp. Quant. Biol.* 20: 225–229. [6]

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. *Anim. Prod.* 8: 95–108. [6]

Robertson, A. 1968. The spectrum of genetic variation. *In* R. C. Lewontin (ed.), *Population biology and evolution*, pp. 5–16. Syracuse University Press, Syracuse, NY [6]

Roff, D. A., and K. Emerson. 2006. Epistasis and dominance: evidence for differential effects in life-history versus morphological traits. *Evolution* 60: 1981–1990. [6]

Roff, D. A., and T. A. Mousseau. 1987. Quantitative genetics and fitness: lessons from *Drosophila. Heredity* 58: 103–118. [6]

Rose, M. 1982. Antagonistic pleiotropy, dominance and genetic variance. *Heredity* 48: 63–78. [6]

Schwaegerle, K. E., and D. A. Levin. 1991. Quantitative genetics of fitness traits in a wild population of phlox. *Evolution* 45: 169–177. [6]

Schwartz, J. 2000. Death of an altruist. *Lingua Franca* 10: 51–61. [6]

Teplitsky, C., J. A. Mills, J. W. Yarrall, and J. Merilä. 2009. Heritability of fitness components in a wild bird population. *Evol.* 63: 716–726. [6]