The Value of Indirect Selection
II. Progeny Testing*

S.R. Searle
Biometrics Unit, Cornell University, Ithaca, N.Y. (USA)

Summary. Conditions are developed under which progeny-testing using indirect selection can give more rapid genetic improvement than using direct selection. Analogous conditions for mass selection are given in Searle [1954].

Key words: Progeny-testing - Indirect Selection - Genetic-improvement

Introduction

The efficiency of indirect selection relative to direct selection, under conditions of mass selection, is considered at length in Searle (1965). Results given there have since been used in studies of egg production, body weight and other traits of the domestic fowl [e.g., Nordskog et al. (1967)] and Festing and Nordskog (1967)] and in studies of milk yield and other traits in dairy cows [e.g., Eisen (1966), Thompson and Loganathan (1968), Wilton and Van Vleck (1968) and Syrstad (1968)]. The relative selection efficiency of mass selection has also been used in fertility studies of Romney ewes [Ch'ang and Rae (1972)] and in simulation studies [Singh et al. (1967)]. In contrast, this paper describes the relative selection efficiency of progeny-testing, and uses it to establish conditions when progeny-testing with indirect selection gives faster genetic improvement than with direct selection.

Relative Selection Efficiency of Progeny-Testing

We consider improvement through selection of a basic trait with genotype represented by y and phenotype by Y. Selecting on Y is called direct selection whereas selecting on something other than Y is called indirect selection. Attention is confined to the additive portion of the genotype and, using the same notation as Searle (1965), heritability in the narrow sense is defined as $h = \sigma_y^2/\sigma_Y^2$, the ratio of the additive genetic variance σ_y^2 to the phenotypic variance σ_Y^2. If R_{Iy} is the correlation between an index I and y, then the relative selection efficiency of indirect selection on I compared to direct mass selection on Y is

$$RSE(I, Y, y) = R_{Iy}/\sqrt{h} . \qquad (1)$$

It is assumed that intensity of selection is the same using I as it is using Y.

We deal with selection indices based on phenotypes of progeny of animals among whom selection is to be practiced. Two common cases are those of using milk yields of dairy cows for selecting among bulls (particularly for use in artificial insemination programs); and of using egg production of domestic fowl for selecting among cockerels. We assume that in the selection process each parent being tested has the same number of progeny, n, and that each progeny has just one record. Although this might sound restrictive, it imposes only minimum limitations in many practical applications; e.g., selection practiced on dairy and poultry sires is largely based upon just first records of their progeny.

Suppose that I_Y is the selection index based on n progeny records on Y. It is well-known [e.g., Falconer (1960)] that the correlation between the progeny-test I_Y and the parent's genotype y is, assuming there is no environmental correlation among progeny records constituting the progeny test,

$$R_{I_y,y} = \sqrt{\frac{nh}{4 + h(n-1)}} \,. \qquad (2)$$

(It is necessary to retain the subscript Y in I_Y to distinguish I_Y from indices I_X, I_{X_1,X_2} and $I_{X,Y}$ that are based on alternative traits X and used subsequently.) Values of (2) are tabulated in Searle (1964), as are values of

$$n = (4/h - 1)/(R_{I_y,y}^{-2} - 1) \,, \qquad (3)$$

obtained by solving (2) for n. This represents the number of progeny required in order that the correlation between I_Y and y shall be of a pre-determined magnitude $R_{I_y,y}$. Naturally, the numbers of progeny required are small for correlations so low as to be valueless (below 0.70 say), but for increasingly larger correlations the numbers increase rapidly. Equations (2) and (3) are, of course, two different ways

of looking at the same expression: (2) provides the correlation between additive genotype and progeny-test for some known number of progeny, whereas (3) is useful for deciding, prior to conducting a progeny-test program, just how many progeny are needed in order to achieve a pre-assigned value of $R_{I_y,y}$.

The correlation in (2) is applicable to any trait; and the corresponding RSE for comparing progeny-testing to performance testing comes from using (2) in (1), giving

$$RSE(I_Y, Y, y) = \sqrt{\frac{n}{4 + h(n-1)}} \,. \qquad (4)$$

However, this expression is inapplicable when performance testing is not possible, namely for sex limited traits such as in sire selection for improvement either in egg-laying in poultry or in milk production in dairy cattle. An alternative to progeny-testing in such cases is to use ancestor records, the efficiency of which is discussed in Searle (1963). In contrast, for traits where performance testing is possible, such as fleece weight in sheep and rate of weight gain in beef cattle, the RSE of (4) is applicable and can be used to answer the question "when is progeny-testing preferred over performance testing". By "preferred

Table 1. Number* of progeny needed so that progeny-testing is at least as good as performance testing (Upper limits $RSE_\infty = 1/\sqrt{h}$ shown in parenthesis)

Heritability (h)										
RSE	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
1.00	5	5	6	6	7	9	11	16	31	(1.00)
1.05	5	6	7	8	9	11	16	30	441	
1.10	6	7	8	9	11	16	27	121	(1.054)	
1.15	6	7	9	11	14	22	59	(1.12)		
1.20	7	8	10	13	18	36	(1.195)			
1.25	8	9	11	15	25	85				
1.40	10	13	18	33	343	(1.29)				
1.50	12	16	26	81	(1.41)					
1.55	13	18	32	222						
1.60	14	20	41	(1.58)						
1.80	19	35	428							
2.00	26	76	(1.83)							
2.20	37	575								
2.50	65	(2.24)								
3.00	351									
3.10	961									
	(3.16)									

* Equation (5): $n = \dfrac{4 - h}{(RSE)^{-2} - h}$.

over" we mean "gives greater rate of genetic improvement", and in this sense progeny-testing will be preferred over performance testing when the RSE of (4) exceeds unity.

Occasions when (4) exceeds 1.0 can be considered in two ways. First, by calculating values of (4) as it stands, for which we observe that (4) is simply (2) multiplied by $h^{-1/2}$. It will be found that except for very small n or for large h, calculated values of (4) usually exceed 1.0; i.e., progeny-testing is usually preferred over performance testing. Furthermore, for each value of h there is an upper limit on (4), obtained by letting n tend to infinity, giving $RSE_\infty = h^{-1/2}$. Note also, that were there to be m records in the performance test of the parent, the RSE of (4) would be lessened through multiplication by $\sqrt{\rho + (1 - \rho)/m}$ where ρ is repeatability. A second way of looking at (4) is to ask the question, "for a given value of h, how many animals are needed in a progeny-test so that the RSE equals some pre-assigned value?" This results in rearranging (4) as

$$n = (4 - h)/[(RSE)^{-2} - h] , \qquad (5)$$

just as (3) was derived from (2). Values of (5) are shown in Table 1 for h " 0.1(0.1)1.0 and for a range of values of RSE $\geq$ 1, representing situations when progeny-testing is as good as or better than performance testing. The upper limit $RSE_\infty = h^{-1/2}$ established from (4) and also arising from the non-negativity of (5) is shown in parentheses at the end of each column of Table 1, indicating that no further values of n exist. The choice of RSE values for the table was made with these upper limits in mind, to have both a range of values and values that are, in most cases, close to the limits.

Progeny-Testing with an Alternative Trait

The relative selection efficiency, under mass selection, of using a trait X rather than Y is

$$p = RSE(X, Y, y) = r\sqrt{h_x/h_y} , \qquad (6)$$

where r is the genotypic correlation between traits X and Y and $h_x = \sigma_x^2/\sigma_X^2$ is the heritability of X in the same way that $h_y = \sigma_y^2/\sigma_Y^2$ is of Y. Considerable

attention is given to p in Searle (1965), and also to RSE(I, Y, y) for I being a linear combination, under mass selection, of either Y and one X, or of two X's. The parameter p also arises in considering $RSE(I_X, Y, y)$ and $RSE(I_X, I_Y, y)$ for progeny-testing indices I_X and I_Y.

Comparisons with performance testing

Denote by I_X the selection index based on single records of n_x progeny. Then, just like (2),

$$R_{I_X,y} = r \sqrt{\frac{n_x h_x}{4 + h_x(n_x - 1)}} = rR_{I_X,x} . \qquad (7)$$

This is the correlation between a progeny-test using X and the additive genotype y of a parent; its values are, as indicated in (7), obtainable by multiplying those of (2) by the genetic correlation r. Also, values of n_x obtained from (7) comparable to (3) are

$$n_x = (4/h_x - 1)/[(R_{I_X,y}/r)^{-2} - 1] ; \qquad (8)$$

i.e., they are the same as (3) using h_x for h and $R_{I_x,y}/r$ in place of $R_{I_y,y}$. Note from (7) and (8) that $R_{I_x,y}$ cannot exceed the genetic correlation - as is eminently reasonable.

The relative selection efficiency of I_X over mass selection on Y is, using (7) in (1),

$$RSE(I_X, Y) = r \sqrt{\frac{h_x}{h_y}} \sqrt{\frac{n_x}{4 + h_x(n_x - 1)}}$$

$$= pRSE(I_X, X, x) \qquad (9)$$

on also using (4) and (6). Just as (4) leads to (5), so does (9) lead to

$$n_x = (4 - h_x)/[(RSE^*)^{-2} - h_x] ,$$

where $RSE^* = RSE(I_X, Y, y)/p$. $\qquad (10)$

Compared to mass selection on Y, n_x of (10) is the number of progeny required for a progeny-test using X to have relative selection efficiency be some pre-assigned value $RSE(I_X, Y, y)$. Since n_x must be positive, (10) implies that the upper limit on RSE* is

$h_x^{-1/2}$ and so the corresponding upper limit on $RSE(I_X, Y, y)$ is $h_x^{-1/2}p$, which equals $rh_y^{-1/2}$ on using (6). Hence, if this upper limit is to exceed unity, r must exceed $\sqrt{h_y}$.

Progeny-tests with the same numbers of progeny

Denote the relative selection efficiency of I_X compared to I_Y by q. Then, in the manner of Searle (1965),

$$q \equiv RSE(I_X, I_Y, y) = \frac{RSE(I_X, Y, y)}{RSE(I_Y, Y, y)} = \frac{pRSE(I_X, X, x)}{RSE(I_Y, Y, y)}$$

$$= p\sqrt{\frac{n_x[4 + h_y(n_y - 1)]}{n_y[4 + h_x(n_x - 1)]}} . \tag{11}$$

(9) and (4) are used in deriving the last two expressions in (11).

Analytical study of the behavior of q for variations in r, h_x, h_y, n_x and n_y is difficult. However, some conclusions can be reached by making the not impractical assumption of having the same number of progeny whether using X or Y, i.e., $n_x = n_y = n$. This gives

$$q = p\sqrt{\frac{4 + h_y(n - 1)}{4 + h_x(n - 1)}} = r\sqrt{\frac{4h_x + h_x h_y(n - 1)}{4h_y + h_x h_y(n - 1)}} , \tag{12}$$

the last of these expressions being that given by Syrstad (1970).

Tabulated values of r for which $q = 1$ for various values of n, h_x and h_y are given by Hinks (1971), together with some conclusions about relative values of p and q. These conclusions and others are now derived analytically from (12). First, progeny-testing with n progeny will be better using the alternative trait X rather than using the basic trait Y when $q > 1$; i.e., from (12), when

$$p > \sqrt{\frac{4 + h_x(n - 1)}{4 + h_y(n - 1)}} \tag{13}$$

or, equivalently,

$$r > \sqrt{\frac{4h_y + h_x h_y(n - 1)}{4h_x + h_x h_y(n - 1)}} . \tag{14}$$

Since $r \leqslant 1$, (14) can be true only for $h_y < h_x$, whereupon from (13) $p > 1$. But from (6), $p > 1$ only when

$r > \sqrt{h_y/h_x}$. Hence, since $q > 1$ only when (14) is satisfied, q can exceed 1.0 only when $h_y < h_x$, $r > \sqrt{h_y/h_x}$ and $p > 1$. However, for $h_y < h_x$ the first expression of (12) shows that $q < p$. Thus we have the following situation: progeny-testing with n progeny will be better using X rather than Y (i.e., $q > 1$) only when $h_y < h_x$ and when (14) is satisfied; and then $p > q > 1$, implying that the increased rate of improvement in y using the alternative trait would be greater under mass selection than under progeny-testing. Note that only if $h_y < h_x$ can it be beneficial to use the alternative trait; and it will be beneficial only if r satisfies (14). Furthermore, the lower limits imposed by (14) on r approach 1.00 as $n \to \infty$. This means that only for small n is there much of a range of values of r that permits $q > 1$ to be satisfied. In addition, $q \to r$ as $n \to \infty$, as is evident from (14).

The inequality $q > 1$ with $n_x = n_y = n$ has been expressed in (14) in terms of limits on r. But it can also be expressed in terms of n as $n < L_1$ for

$$L_1 \equiv L_1(h_x, h_y, r) = 1 + \frac{4(r^2 h_x - h_y)}{h_x h_y(1 - r^2)}$$

$$= 1 + \frac{4(p^2 - 1)}{h_x(1 - r^2)} . \tag{15}$$

[The functional form $L_1(h_x, h_y, r)$ is introduced for use on page 294]. Clearly L_1 is positive for $p > 1$, so that when (13) and (14) are satisfied so also is $n < L_1$. Thus is established an upper limit L_1 such that progeny-testing using an alternative trait is better than using the basic trait with the same number of progeny. This means that for an alternative trait with $p > 1$, progeny-testing based on that trait can be better than using the basic trait with the same number of progeny only up to a limited number of progeny - and beyond that limit progeny-testing using the basic trait is better. Note that the limit exists only when $p > 1$, which requires $r > \sqrt{h_y}$ and $h_x > h_y/r^2$, as discussed in Searle (1965).

Progeny-tests with different numbers of progeny

We now consider progeny-tests based on different numbers of progeny, using q of (11). We have just

seen when $p > 1$ that for $n < L_1$, progeny-testing with $n_x = n_y = n$ progeny is better using the alternative trait than using the basic trait, i.e., $q < 1$. For some cases of $n_x = n_y < L_1$, the value of q will exceed 1.00 sufficiently so as to enable use of $n_x < n_y$ and still have $q > 1$. This means that for some $n_y < L_1$, using fewer than n_y progeny on the alternative trait (i.e., $n_x < n_y$) will be better than using n_y progeny with the basic trait. For example, with $r = 0.9$, $h_y = 0.2$ and $h_x = 0.6$ equation (15) gives $L_1 = 51$; and for $n_x = n_y = n = 20 < L_1 = 51$, equation (12) gives $q = 1.11$. This exceeds 1.00 sufficiently that we can in fact use the alternative trait with $n_x < 20$ and from (11) still have $q > 1$, i.e., still have the alternative trait being better than the basic trait. Thus for $n_x = 15$, equation (11) gives $q = 1.07$ and for $n_x = 12$, $q = 1.04$, so that either 15 or 12 progeny using the alternative trait are better than 20 progeny using the basic trait. Clearly, 15 or 12 progeny on the alternative trait ($q = 1.07$ and $q = 1.04$, respectively,) are not as good as 20 progeny are ($q = 1.11$), but the important thing is that in these cases q still exceeds unity but with $n_x < n_y$.

Suppose we ask quite generally "how many progeny are needed using X to be equivalent to progeny-testing with n_y progeny using Y?" The answer is the solution for n_x to the equation $q = 1$. Using (11), this is $n_x = \tilde{n}_x$ for

$$\tilde{n}_x = \tilde{n}_x(x_y, h_x, h_y, r) = \frac{n_y h_y (4 - h_x)}{h_x [r^2 (4 - h_y) - n_y h_y (1 - r^2)]} . \tag{16}$$

Since $n_x > 0$ we immediately see that this equivalence can occur only if $n_y < L_2$ for

$$L_2 \equiv L_2(h_y, r) = \frac{r^2 (4 - h_y)}{h_y (1 - r^2)} . \tag{17}$$

[Again the functional forms are introduced for subsequent use.] Since we have already seen that this kind of equivalence occurs for $n_x = n_y = n < L_1$ of (15), we would expect L_2 of (17) to exceed L_1; and it does, as is easily shown. Thus L_2 is an upper limit to the number of progeny using the basic trait, above which not even an infinite number of progeny using the alternative trait can yield equivalent progeny-testing proce-

dures. For example, with $r = 0.9$, $h_y = 0.2$, the limit given by (17) is $L_2 = 81$. Hence progeny-testing using alternative traits with genetic correlation 0.9 can never be preferred to using the basic trait with 81 or more progeny, no matter how many progeny are available for the alternative trait, nor what the heritability of that alternative trait may be. For example, with $h_x = 0.6$, $h_y = 0.9$, $r = 0.9$ and with $n_y = 100 > L_2 = 81$ and $n_x = 200$, equation (11) yields $q = 0.97 < 1$. Thus L_1 and L_2 divide values of n_y into three groups:

Group I: for $n_y < L_1$, there exists $n_x < n_y$ such that I_X is preferred to I_Y.

Group II: for $L_1 \leqslant n_y < L_2$, there exists $n_x > n_y > L_1$ with I_X preferred to I_Y.

Group III: for $n_y \geqslant L_2$, I_X is never preferred to I_Y.

The example that has been used is illustrated in Fig. 1, which shows $R_{I_y, y}$ and $R_{I_x, y}$ plotted against n. The division of n-values into three groups separated by $L_1 = 51$ and $L_2 = 81$ is self-evident.

The boundary between Group I and Group II values of n_y is L_1 of (15) and that between Groups II and III is L_2 of (17). For Group II values of n_y, where $n_x > n_y$, the lack of complete correlation between the alternative and basic traits (i.e., $r < 1$) can be compensated for in using the alternative trait by having more progeny, but only to a certain extent. There is a definite limit beyond which this compensatory effect cannot be had, and that limit is L_2 which represents the number of progeny using the basic trait beyond which not any number of progeny using the alternative trait can be equivalent. As is evident from (17), values of L_2 are smaller for large values of h_y than for small. For example, when the basic
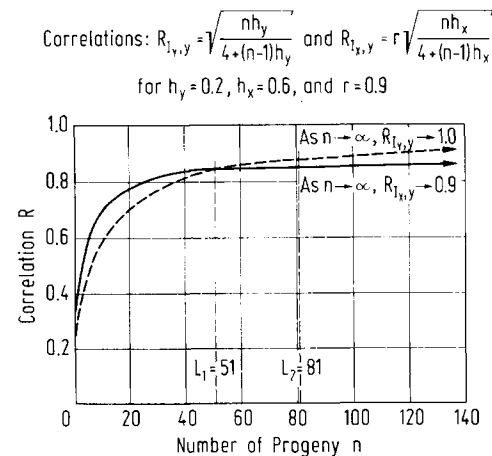

Correlations: $R_{I_y, y} = \sqrt{\dfrac{n h_y}{4 + (n-1) h_y}}$ and $R_{I_x, y} = r \sqrt{\dfrac{n h_x}{4 + (n-1) h_x}}$

for $h_y = 0.2$, $h_x = 0.6$, and $r = 0.9$

Fig. 1

trait has heritability of 0.60 or more, L_2 is less than 20 unless r is 0.90 or more. This simply means that in these situations using an alternative trait can be equivalent to using a basic trait having high heritability only when the latter is used on relatively few progeny. When the basic trait is used with more than 20 or so progeny, it is then more reliable than any progeny-test using an alternative trait could be, no matter how many progeny were available. It is also clear from (17) that values of L_2 are larger for large values of r than for small, particularly when r is close to 1.00, corresponding to the upper limit on n_y being infinite when r = 1. In this case the high genetic correlation between the alternative and basic traits means that the two traits are sufficiently correlated genetically that a progeny-test using the alternative trait can be equivalent to one using the basic trait even for large numbers of progeny in the latter. When they are perfectly correlated (r = 1), the limit is infinite, meaning that for all progeny-tests using the basic trait equivalent tests using the alternative trait can be found, in which case $n_x = n_y (4/h_x - 1) / (4/h_y - 1)$. A final and obvious comment, but one worthy of note, is that L_2 is independent of h_x, the heritability of the alternative trait. Values of equations (14) through (17) are tabulated and discussed in an early version of this paper, available from the author.

Using Two Alternative Traits

Suppose we wish to compare progeny-tests I_1 and I_2 using two different alternative traits, X_1 and X_2 respectively, for selection for improvement in y. Let h_1, h_2 be the heritabilities of the alternative traits; r_1, r_2 their genetic correlations with the basic trait and n_1, n_2 the number of progeny used, respectively; and let the genetic and phenotypic correlations between the two traits be r_{12} and R_{12}, respectively. This is the notation of Searle (1965), illustrated there in Fig.3.

Comparing Two Alternatives

The relative selection efficiency of I_1 compared to I_2 is

$$RSE(I_1, I_2, y) = \frac{R_{I_1,y}}{R_{I_2,y}} =$$

$$= \frac{r_1}{r_2} \sqrt{\frac{n_1 h_1}{4 + h_1(n_1 - 1)} \Big/ \frac{n_2 h_2}{4 + h_2(n_2 - 1)}} = \frac{r_1 R_{I_1,x_1}}{r_2 R_{I_2,x_2}} . \quad (18)$$

Progeny-testing with I_1 will be equivalent to I_2 when $RSE(I_1, I_2, y) = 1$. Equating (18) to unity and solving for n_1, the solution is $n_1 = \tilde{n}_1$ for

$$\tilde{n}_1 = \frac{n_2 r_2^2 h_2 (4 - h_1)}{h_1 [r_1^2 (4 - h_2) - n_2 h_2 (r_2^2 - r_1^2)]} , \quad (19)$$

a result similar to $\tilde{n}_x$ of (16). In fact, on defining $\rho_1^2 \equiv r_1^2/r_2^2$ for $r_1^2 < r_2^2$, we can express (19) as $\tilde{n}_1 = \tilde{n}_x(n_2, h_1, h_2, \rho_1)$. Applying the arguments of Section 3.3 leads to the following conclusions.

Case A: $r_1^2 < r_2^2$, with $\rho_1^2 \equiv r_1^2/r_2^2 < 1$

I: For $n_2 < L_1(h_1, h_2, \rho_1)$ there exists $n_1 < n_2$ such that I_1 is preferred to I_2.
II: For $L_1(h_1, h_2, \rho_1) \leq n_2 < L_2(h_2, \rho_1)$ there exists $n_1 > n_2 > L_1(h_1, h_2, \rho_1)$ such that I_1 is preferred to I_2.
III: For $n_2 \geq L_2(h_2, \rho_1)$, I_1 is never preferred to I_2.

We label the preceding conclusions Case A because (19) must also be considered for $r_1^2 = r_2^2$ and for $r_1^2 > r_2^2$.

Case B: $r_1^2 = r_2^2$ Equation (19) reduces to $\tilde{n}_1 = n_2(4/h_1 - 1)/(4/h_2 - 1)$ and we have the following.

I: For any h_1 and h_2 there always exists $n_1 > \tilde{n}_1$ such that I_1 is preferred to I_2; $h_1 = h_2$ implies $\tilde{n}_1 = n_2$, and $h_1 < h_2$ implies $\tilde{n}_1 > n_2$.
II: For $h_1 > h_2$, $\tilde{n}_1 < n_2$ and there exists $n_1 < n_2$, i.e. $\tilde{n}_1 < n_1 < n_2$ such that I_1 is preferred to I_2.

Case C: $r_1^2 > r_2^2$, with $\rho_2^2 \equiv r_2^2/r_1^2 < 1$. Consider the difference $\tilde{n}_1 - n_2$, which from (19) is

$$\tilde{n}_1 - n_2 = \frac{n_2}{h_1} \left[\frac{4(r_2^2 h^2 - r_1^2 h_1) + (n_2 - 1) h_1 h_2 (r_2^2 - r_1^2)}{r_1^2 (4 - h_2) - n_2 h_2 (r_2^2 - r_1^2)} \right] . \quad (20)$$

The denominator is always positive and the second term of the numerator is negative. The following conclusions ensue.

I: For $r_1^2 h_1 \geq r_2^2 h_2$, $\tilde{n}_1 < n_2$ and there exists $n_1 < n_2$, i.e., $\tilde{n}_1 < n_1 < n_2$, such that I_1 is preferred to I_2.

II: For $r_1^2 h_1 < r_2^2 h_2$, $\tilde{n}_1 > n_2$ if $r_1^2 h_1$ is sufficiently less than $r_2^2 h_2$ to make the numerator of (20) positive. This leads to $n_2 < L_1(h_2, h_1, \rho_2)$ being the condition for I_1 to be preferred over I_2 for $n_1 > > n_2$, i.e. for $n_1 > \tilde{n}_1 > n_2$.

III: Conversely, with $r_1^2 h_1 < r_2^2 h_2$ and $n_2 \geq L_1(h_2, h_1, \rho_2)$ then I_1 can be preferred over I_2 for $n_1 < n_2$, i.e. for $\tilde{n}_1 < n_1 < n_2$.

Combining Two Alternatives

The index combining the progeny-tests of the two traits can be taken as

$$I_{12} = b_1 \overline{X}_1 + b_2 \overline{X}_2$$

where $\overline{X}_1$ and $\overline{X}_2$ are the means of single records of traits X_1 and X_2 on the same n progeny. Then the correlation R_{Iy} is well known to be [e.g. Searle (1963)]

$$R_{I_{12}y}^2 = \underset{\sim}{w'} \, \underset{\sim}{V}^{-1} \underset{\sim}{w}/h$$

Then the relative selection efficiency of I_{12} compared to I_Y is

$$RSE(I_{12}, I_Y, y) = \frac{R_{I_{12},y}}{\sqrt{nh/[4 + (n-1)h]}}$$

$$= \sqrt{\frac{\left[1 + \frac{1}{4}(n-1)h\right] R_{I_{12},y}^2}{\frac{1}{4}nh}} \, . \tag{22}$$

The intractability of (22) with respect to n, h_1, h_2, r_1, r_2, r_{12} and R_{12} appears to preclude establishing any simple or workable conditions on these parameters such that $RSE(I_{12}, I_Y, y) \geq 1$; i.e., conditions under which I_{12} is to be preferred to I_Y cannot be established other than requiring that (22), with (21) substituted therein, exceed unity.

Combining an Alternative Trait With the Basic Trait

Instead of selecting on I_{12}, one might use I_{1Y}, the index made up of Y and one of the alternative traits, X_1 say. In that case, h_2, r_2, r_{12} and R_{12} of (21) become h, 1, r and R respectively and (21) reduces to

$$R_{I_{1y},y}^2 = \frac{\frac{1}{4}n\left[\frac{1}{4}(n-1)hh_1(1-r^2) - (1-r^2)h_1 + h_1 + h - 2rR\sqrt{hh_1}\right]}{1 - R^2 + (1/16)(n-1)^2(1-r^2)hh_1 + \frac{1}{4}(n-1)(h_1 + h - 2rR\sqrt{gg_1})} \, . \tag{23}$$

where

$$\underset{\sim}{w} = \begin{bmatrix} cov(\overline{X}_1 y) \\ cov(\overline{X}_2 y) \end{bmatrix}, \quad \text{and} \quad \underset{\sim}{V} = var\begin{bmatrix} \overline{X}_1 \\ \overline{X}_2 \end{bmatrix},$$

$\underset{\sim}{w}$ being the vector of covariances of y with $\overline{X}_1$ and $\underset{\sim}{V}$ the variance-covariance matrix of $\overline{X}_1$ and $\overline{X}_2$. It is readily shown that

$$R_{I_{12},y}^2 = \frac{1}{h}\begin{bmatrix} \frac{1}{2}r_1\sqrt{hh_1} & \frac{1}{2}r_2\sqrt{hh_2} \end{bmatrix} \times$$

$$\times \begin{bmatrix} \dfrac{1 + \frac{1}{4}(n-1)h_1}{n} & \dfrac{R_{12} + \frac{1}{4}(n-1)r_{12}\sqrt{h_1 h_2}}{n} \\[2ex] \dfrac{R_{12} + \frac{1}{4}(n-1)r_{12}\sqrt{h_1 h_2}}{n} & \dfrac{1 + \frac{1}{4}(n-1)h_2}{n} \end{bmatrix}^{-1}$$

It can then be shown, using (23) in (22), that the numerator of $[RSE(I_{1Y}, I_Y, y)]^2 - 1$ is $(r\sqrt{h_1} - R\sqrt{h})^2$ which is positive; i.e. $RSE(I_{1Y}, I_Y, y) > 1$, meaning that selecting on an index made up of Y and an alternative is always better than selecting on one based on Y alone - a not unexpected result.

Sampling Variances

Preferences for alternative traits have been discussed entirely in terms of genetic parameters, which have to be estimated. Decisions about using alternative traits therefore have to be made on the basis of estimated RSE's. Even though the sampling distribution

$$\times \begin{bmatrix} \frac{1}{2}r_1\sqrt{hh_1} \\ \frac{1}{2}r_2\sqrt{hh_2} \end{bmatrix} = \frac{\frac{1}{4}n\left[\frac{1}{4}(n-1)h_1 h_2(r_1^2 + r_2^2 - 2r_1 r_2 r_{12}) + r_1^2 h_1 + r_2^2 h_2 - 2r_1 r_2 R_{12}\sqrt{h_1 h_2}\right]}{1 - R_{12}^2 + \frac{1}{16}(n-1)^2(1-r_{12}^2)h_1 h_2 + \frac{1}{4}(n-1)(h_1 + h_2 - 2r_{12}R_{12}\sqrt{h_1 h_2})} \, . \tag{21}$$

of these estimates is unknown, there would be some comfort to using the estimates if we could have even approximate standard errors for them. This has been done in the case of mass selection, where p of (6) is the appropriate RSE. Approximate sampling variances of an estimator of p have been obtained by Searle (1965), based on parent-progeny records for estimating genetic parameters, and by Scheinberg (1967) based on sibship data. In both cases, derivation of the sampling variance is tedious, is approximate only, and results in an expression that is too complicated to be studied analytically. Only numerical studies are possible. The complications are aggravated in the case of the sampling variance of the relative selection efficiency of progeny-testing, of say $RSE(I_y, Y, y)$ of (4) or of q of (11), because whereas p is just a simple product of r, $h_x^{1/2}$ and $h_y^{-1/2}$, q involves $[4 + h_x(n_x - 1)]^{-1/2}$ and $[4 + h_y(n_y - 1)]^{1/2}$; and the methods of deriving sampling variances do not lend themselves at all easily to handling this kind of expression. For example, on adapting Scheinberg's (1967) notation slightly, his expression (8) is $\hat{p} = \hat{G}_{xy}\hat{P}_{yy}^{1/2}\hat{P}_{xx}^{-1/2}\hat{G}_{yy}^{-1}$ where $\hat{P}_{yy}$ (and $\hat{G}_{yy}$) are estimated phenotypic (and genotypic) variances, and $\hat{G}_{xy}$ is an estimated genotypic covariance. The comparable expression for $\hat{q}$ is $\hat{q} =$

$= (\sqrt{n_1/n_y})\hat{G}_{yy}\hat{H}_{yy}^{1/2}\hat{H}_{xx}^{-1/2}\hat{G}_{yy}^{-1}$ for $H_{yy} = 4P_{yy} + (n_y - 1)G_{yy}$. Attempts at following Scheinberg's procedure for deriving $var(\hat{q})$ yield expressions vastly more complicated than his, including the added complexity of involving n_x and n_y. Their practical value would therefore appear to be very limited.

A feasible alternative is to obtain $var(\hat{q}) \doteq \hat{\lambda}^2 var(\hat{p})$ directly from (11) for

$$\hat{\lambda}^2 = \frac{n_x[4 + \hat{h}_y(n_y - 1)]}{n_y[4 + \hat{h}_x(n_x - 1)]} .$$

Although not an exact relationship, this is the conditional variance of $\hat{q}$ given $\hat{\lambda}^2$. Cautious use of this in the easily computed form $\widehat{var}(\hat{q}) \doteq \hat{\lambda}^2 \widehat{var}(\hat{p})$, using $var(\hat{p})$ from (11) of Scheinberg (1967), might be little worse than the excessive complexity of the preceding paragraph.

Literature

Ch'ang, T.S.; Rae, A.L.: The genetic basis of growth, reproduction and maternal environment in Romney ewes. II. Genetic covariance between hogget characters, fertility and maternal environment of the ewe. Austr. J. Agric. Res. 23, 149-165 (1972)

Eisen, E.J.: Effect of the biometrical relationship among total milk yield, milk constituent and percent of milk constituent on response to selection. J. Dairy Sci. 49, 1230-1234 (1966)

Falconer, D.S.: Introduction to Quantitative Genetics. Edinburgh: Oliver and Boyd 1960

Festing, M.; Nordskog, A.W.: Response to selection for body weight and egg weight in chickens. Genetics 55, 219-231 (1967)

Hinks, C.J.M.: A note on the relative magnitude of genetic responses to direct and indirect selection. Acta Agric. Scand. 21, 264-266 (1971)

Nordskog, A.W.; Festing, M.; Verghese, M.W.: Selection for egg production and correlated responses in the fowl. Genetics 55, 179-191 (1967)

Scheinberg, E.: The sampling variance of the relative efficiency of indirect to direct selection when using variance-covariance components. Austral. J. Statistics 9, 35-40 (1967)

Searle, S.R.: The efficiency of ancestor records in animal selection. Heredity 18, 351-360 (1963)

Searle, S.R.: Progeny-tests of sire and son. J. Dairy Sci. 47, 414-420 (1964)

Searle, S.R.: The value of indirect selection: I. Mass selection. Biometrics 21, 682-707 (1965)

Singh, R.K.; Bellman, K.; Ahrens, H.: Die Abhängigheit der korrelierten Antwort auf die Selektion von Kopplung und Heritabilität. Biometr. Z. 9, 240-249 (1967)

Syrstad, Ola: Studies on dairy herd records. V. Direct and correlated response to selection. Acta Agric. Scand. 18, 87-90 (1968)

Syrstad, Ola: Estimating direct and correlated response to selection. A note for clarification. Acta Agric. Scand. 20, 205-206 (1970)

Thompson, N.R.; Loganathan, S.: Composition of cows' milk. II. Genetic influences. J. Dairy Sci. 51, 1933-1935 (1968)

Wilton, J.W.; Van Vleck, L.D.: Selection of dairy cows for economic merit. J. Dairy Sci. 51, 1680-1688 (1968)

S.R. Searle
Biometrics Unit,
Cornell University,
Ithaca, New York 14853
(USA)