

# Mass selection in livestock using limited testing facilities

L Ollivier

*Institut National de la Recherche Agronomique, Station de Génétique Quantitative et Appliquée, 78350 Jouy-en-Josas, France*

(Received 29 May 1989; accepted 20 September 1989)

**Summary** – This paper considers the problem of maximizing the expected annual response to mass selection when testing facilities are limited and so do not allow testing of all potential candidates. In such situations, there is room for variation both in the proportion of breeding animals selected on the basis of the test results, and in the allocation of testing places between male and female candidates. When testing facilities are very limited (case 1), males have priority in testing and the maximum proportion to select based on test results is 27%. This means that it is then better to use untested males, *i.e.* taken at random, than males which are in the lower 73%. This situation holds until the ratio ( $k$ ) of tested to potential candidates reaches  $k_1 = 1.85/c(4a\lambda + 1)$ , where  $c$  is the degree of polygyny (mating ratio),  $a$  the age at first offspring ( $yr$ ) and  $\lambda$ , the annual fecundity (*i.e.* half the dam progeny crop). As  $k$  increases above  $k_1$  (case 2), all replacement males should be tested and testing space should be entirely devoted to males, with random choice of females. This situation holds until  $k$  reaches a critical value,  $k_2$ , above which testing space should be equally distributed between the 2 sexes (case 3). The value of  $k_2$ , obtained iteratively for any given set of parameters  $c$ ,  $a$  and  $\lambda$ , as defined above, is shown to increase when  $c$  increases and when  $a\lambda$  decreases. The strategies recommended, which imply contrasting turn-over rates between selected candidates and candidates chosen at random, are compared to those aimed at maximizing selection intensity for a fixed value of the generation interval. Numerical examples are provided, covering the range of situations prevailing in farm livestock species.

**mass selection / selection response / selection intensity / generation interval**

**Résumé** – La sélection massale chez les animaux domestiques avec une capacité de contrôle limitée – Cet article traite de la maximisation du gain génétique annuel attendu en sélection massale quand la capacité de contrôle est limitée et ne permet pas de contrôler tous les candidats potentiels à la sélection. Dans une telle situation, on peut faire varier à la fois la proportion des reproducteurs sélectionnés sur leur résultat de contrôle et la répartition des places de contrôle entre les 2 sexes. Quand la capacité de contrôle est restreinte (cas 1), les mâles ont la priorité et le taux de sélection maximal à l'issue des contrôles est de 27%. Il vaut mieux alors utiliser des mâles non contrôlés, *c'est-à-dire* choisis au hasard, que des mâles se trouvant dans les 73% inférieurs. Cette situation prévaut tant que le rapport ( $k$ ) des candidats contrôlés aux candidats potentiels ne dépasse pas  $k_1 = 1,85/c(4a\lambda + 1)$ , où  $c$  est le degré de polygynie (nombre de reproducteurs femelles/nombre de reproducteurs mâles), à l'âge au 1<sup>er</sup> descendant ( $an$ ) et  $\lambda$ , la fécondité annuelle (*c'est-à-dire* la moitié du nombre de descendants produits annuellement par

femelle). Quand  $k$  dépasse  $k_1$  (cas 2) tous les mâles de renouvellement doivent être contrôlés et toutes les places de contrôle doivent être réservées aux mâles, les femelles étant choisies au hasard. Cette situation prévaut jusqu'à une valeur critique  $k = k_2$ , au-dessus de laquelle les places doivent être également réparties entre les 2 sexes (cas 3). On montre que cette valeur  $k_2$ , qui est obtenue par itération pour tout ensemble donné des paramètres  $c$ ,  $a$  et  $\lambda$ , définis ci-dessus, augmente avec  $c$  et diminue quand  $a\lambda$  augmente. Les stratégies recommandées, qui impliquent des taux de renouvellement très différents entre les candidats sélectionnés et les candidats choisis au hasard, sont comparées à celles qui visent à maximiser l'intensité de sélection à intervalle de génération fixé. Des exemples sont donnés pour illustrer le cas des diverses espèces animales domestiques.

sélection massale / réponse à la sélection / intensité de sélection / intervalle de génération

## INTRODUCTION

Mass selection is a simple and widely used selection method for farm animals. Considering a trait expressed in both sexes, and following a normal distribution, the expected annual response can be shown to be a function of the mean ages of males and females at culling. The maximum response is obtained by determining an optimal balance between selection intensities and generation intervals, as shown by Ollivier (1974) for the case when all potential candidates are tested. The purpose of this paper is to extend the treatment to situations where testing facilities are limited and so do not allow testing of all potential candidates. In such cases, there is room for variation both in the proportion of breeding animals selected on the basis of the test results, and in the allocation of testing places between male and female candidates. The effect of such a variation on the overall selection intensity has previously been considered by Smith (1969).

### *The general method*

Dickerson and Hazel (1944) gave a general formula for the expected annual response to selection,  $R_a = (i_1 + i_2)/(t_1 + t_2)$ , as a function of female and male selection intensities ( $i_1$  and  $i_2$ , respectively) and generation intervals ( $t_1$ ,  $t_2$ ),  $R_a$  being expressed in genetic standard deviations for a trait assumed to have a heritability equal to 1. With selection of respective proportions  $f$  and  $m$  of the females and males required for breeding, and corresponding proportions  $1 - f$  and  $1 - m$  taken at random, the expected annual response becomes:

$$R_a = (fi_1 + mi_2)/(ft_{11} + (1 - f)t_{12} + mt_{21} + (1 - m)t_{22}) \quad (1)$$

where  $t_{11}$  and  $t_{12}$  are the generation intervals for the females selected and the females taken at random, respectively, and  $t_{21}$  and  $t_{22}$  are similarly defined for males.

If selection is by truncation of a normal distribution,  $i_1 = z_1 n_1$ , where  $n_1$  is the number of female candidates tested per female selected and  $z_1$  the ordinate of the normal curve for a proportion  $1/n_1$  selected, and  $i_2$  is similarly defined. Moreover, generation intervals may be expressed as functions of demographic parameters pertaining to any given species, and of the distribution of testing space between males and females. Using the simple demographic model assumed by Ollivier (1974), for instance, one can write:

$$t_{11} = a + fn_1/2l_1 \quad (2a)$$

$$t_{12} = a + (1 - f)/2(\lambda - l_1) \quad (2b)$$

$$t_{21} = a + mn_2/2cl_2 \quad (2c)$$

$$t_{22} = a + (1 - m)/2c(\lambda - l_2) \quad (2d)$$

where  $a$  is the parents' age (in years) at birth of first offspring, assumed equal for both sexes;  $c$  is the degree of polygyny, or mating ratio;  $\lambda$  is the annual female fecundity, referring to the number of candidates of 1 sex (sex ratio assumed to be 1/2) able to breed successfully;  $l_1$  and  $l_2$  are the respective numbers of female and male candidates tested annually per dam.

Expressions (2) are based on the definition taken for the generation interval, which is assumed to be the arithmetic mean of the parents' ages at birth of first ( $a$ ) and of last offspring. The latter is determined by the time necessary to replace 1 breeding animal, either selected among  $n$  candidates or taken at random. For instance, knowing that  $l_1$  female candidates are tested annually per breeding female, *ie*,  $l_1/f$  candidates per female selected, and that each selected female is chosen among  $n_1$  candidates, the time required is  $fn_1/l_1$  years, which leads to eqn(2a). On the other hand,  $(\lambda - l_1)$  females are untested, *ie*,  $(\lambda - l_1)/(1 - f)$  per female chosen at random. The time necessary to obtain 1 candidate, if one takes the first born, is  $(1 - f)/(\lambda - l_1)$ , which leads to eqn(2b). Equations (2c) and (2d) are similarly obtained.

Now  $l_1$  and  $l_2$  depend on the overall testing capacity, defined as the proportion  $k$  of available candidates which can be tested annually, and of the distribution of testing places between females and males, defined by the sex ratio  $\alpha$  among the tested candidates, so that:

$$l_1 = 2k\lambda(1 - \alpha) \text{ and } l_2 = 2k\lambda\alpha. \quad (3)$$

The possible range of  $\alpha$  extends from 0 to 1 as long as  $k \leq 0.5$ . Then, as  $k$  exceeds 0.5, the range is progressively narrowed, until  $\alpha = 0.5$  when  $k = 1$ .

**Case 1: only males are tested ( $\alpha = 1$ ); a proportion ( $m < 1$ ) of males required for breeding is tested**

In this case,  $f = l_1 = 0$  and  $l_2 = 2k\lambda$ . Expression (1) reduces to a function of 2 variables,  $m$  and  $n_2$ , such that:

$$R_a = mz_2n_2/(t_{12} + mt_{21} + (1 - m)t_{22}) \quad (4a)$$

with

$$t_{12} = a + 1/2\lambda, t_{21} = a + mn_2/4ck\lambda \text{ and } t_{22} = a + (1 - m)/2c\lambda(1 - 2k) \quad (4b)$$

The maximum of  $R_a$ , with respect to  $m$  is obtained for:

$$m^2 = (1 + (1 - 2k)c(4a\lambda + 1))/(1 + (1 - 2k)n_2/2k) \quad (5)$$

With this value of  $m$ ,  $R_a$  becomes a quantity approximately proportional to  $z_2n_2^{0.5}$ , which is maximum for  $n_2 \simeq 3.7$ . Thus, the critical value of  $k$  for which  $m = 1$ , is from eqn(5):

$$k_1 = n_2/2c(4a\lambda + 1) \quad (6a)$$

or, with  $n_2 \simeq 3.7$ ,

$$k_1 \simeq 1.85/c(4a\lambda + 1) \quad (6b)$$

Consequently, when testing capacity is limited to a value  $k < k_1$ , a proportion of untested males should be used, in order to maintain a constant proportion selected of about 27% (1/3.7) among those tested. Under these conditions, the expected annual response is approximately proportional to  $k^{0.5}$ , as

$$R_a \simeq 0.9\lambda(ck/(4a\lambda + 1))^{0.5} \quad (7)$$

**Case 2: only males are tested ( $\alpha = 1$ ); all males required for breeding are tested ( $m = 1$ )**

As  $k$  becomes equal to  $k_1$ , and then increases above  $k_1$ ,  $m = 1$  and eqn(4a) reduces to:

$$R_a = z_2 n_2 / (t_{12} + t_{21}) \quad (8)$$

which can be maximized iteratively with respect to  $n_2$ . But the question then arises as to whether a higher response can be expected by diverting some testing space for the selection of females. This case will now be considered.

**Case 3: all males tested ( $m = 1$ ) and a proportion of females ( $\alpha < 1$ ;  $f > 0$ )**

With selection of all males ( $m = 1$ ), and of a proportion ( $f$ ) of the females required for breeding,  $R_a$  becomes:

$$R_a = (f z_1 n_1 + z_2 n_2) / (f t_{11} + (1 - f) t_{12} + t_{21}) \quad (9)$$

which is a function of  $f$ ,  $\alpha$ ,  $n_1$  and  $n_2$  for any given testing capacity.

It can easily be shown that the derivative of  $R_a$ , with respect to  $f$ , is positive when  $0 < f < 1$ , provided  $2i_2 > i_1$ . As selection should generally be more intense in males ( $i_2 > i_1$ ), this condition is always fulfilled, and the optimum value of  $f$  is therefore 1, irrespective of the other parameters.

Then, assuming  $f = 1$  (ie, all females required for breeding are tested), the question is how to allocate the testing places between 2 sexes, within the limits previously indicated for the sex ratio  $\alpha$  among tested candidates. In fact, the value of  $R_a$  is rather insensitive to variations of  $\alpha$  (although the optimal value of  $\alpha$  is slightly below 0.5), as shown by Ollivier (1988: see eqn(6), p 446). One can then take  $\alpha$  to be 0.5, and the optimal values of  $n_1$  and  $n_2$  are obtained by maximizing:

$$R_a = (z_1 n_1 + z_2 n_2) / (t_{11} + t_{21}) \quad (10)$$

where  $t_{11} = a + n_1/2k\lambda$ , and  $t_{21} = a + n_2/2ck\lambda$ , as  $l_1 = l_2 = k\lambda$ .

For any given testing capacity, the maximum of eqn(10) can be compared to the maximum of eqn(8) considered in case 2, and (by iteration) the  $k_2$  value yielding equal responses in the 2 cases is obtained. Thus, when testing capacity is below  $k_2$ , all testing space should be devoted to males, and when  $k > k_2$ , it should be equally distributed between the 2 sexes.

The strategies to be applied in each of the 3 cases considered are summarised in Table I.

**Table I.** Summary of mass selection strategy as a function of testing capacity ( $k$ ).

	Value of $k$		
	From 0 to $k_1$ (case 1)	From $k_1$ to $k_2$ (case 2)	From $k_2$ to 1 (case 3)
Optimal proportion selected among those tested			
females	none tested	none tested	$1/n_1$
males	$1/3.7$	$1/n_2$	$1/n_2$
Optimal generation interval			
females selected ( $t_{11}$ )	-	-	$a + n_1/2k\lambda$
females taken at random ( $t_{12}$ )	$a + 1/2\lambda$	$a + 1/2\lambda$	-
males selected ( $t_{21}$ )	$a + (2a + 1/2\lambda)(k_1/k)^{0.5}$	$a + n_2/4ck\lambda$	$a + n_2/2ck\lambda$
males taken at random ( $t_{22}$ )	$a + (1 - (k_1/k)^{0.5})/2c\lambda$	-	-
Equation to be maximized	(4)	(8)	(10)

$k$  = ratio of individuals tested to those available for testing.

$c$  = degree of polygyny (mating ratio).

$a$  = age of parents at birth of first offspring.

$\lambda$  = annual female fecundity.

$n_1(n_2)$  = number of candidates tested per female (male) retained, which maximizes expected response from equations (8) or (10) given in text.

**Numerical illustration**

As an illustration of the above results, Table II gives  $k_1$  and  $k_2$  values for 9 sets of demographic parameters implying 3 values of  $a\lambda$  (0.5, 1 and 5) valid for sheep, cattle and pigs, respectively, and 3 degrees of polygyny, either corresponding to natural mating ( $c = 10$ ) or artificial insemination ( $c = 100$  or  $1000$ ). The Table also gives the expected response for  $k = k_1$  and  $k = k_2$ , expressed relative to the maximum response expected with  $k = 1$ .

The Table clearly shows that, for a given degree of polygyny,  $k_1$  and  $k_2$  both decrease when fecundity increases. For species of high fecundity, such as poultry and rabbits,  $k_1$  becomes negligible and the low value of  $k_2$  is likely to fall below the actual testing capacity, owing to the low cost of testing. Therefore, case 3 will usually apply to those species. On the other hand,  $k_1$  decreases when polygyny increases, as it is inversely proportional to  $c$ , (from eqn(6)) whereas,  $k_2$  increases with  $c$  up to a point where, particularly when fecundity is low, a large proportion of the maximum response can be expected from testing males only. It is also worth noting that when fecundity is low (below a limit which is somewhere between 1 and 5 for  $a\lambda$ ), the critical testing capacity,  $k_2$ , is above 0.5. As this corresponds to situations when all males are tested, it means that the expected response remains constant, and above the maximum of eqn(10), for  $0.5 \leq k \leq k_2$ . The evolution of the maximum annual response, as a function of testing capacity, therefore follows

**Table II.** Critical testing capacities in mass selection.

Value of $c$	10			100			1000		
Value of $a\lambda$	0.5	1	5	0.5	1	5	0.5	1	5
$k_1$	0.061	0.037	0.0088	0.0061	0.0037	0.00088	0.0006	0.0004	0.00009
$R_a(k_1)/R_a(1)$	0.39	0.33	0.22	0.27	0.24	0.18	0.21	0.20	0.16
$k_2$	0.77	0.65	0.28	0.87	0.75	0.35	0.94	0.81	0.41
$R_a(k_2)/R_a(1)$	0.85	0.81	0.66	0.93	0.89	0.76	0.96	0.93	0.82

$k_1$  = testing capacity below which a proportion of untested males should be used.

$k_2$  = testing capacity below which only males should be tested.

$c$ ,  $a$  and  $\lambda$  are as defined in Table I.

$R_a(k)$  = expected annual response for testing capacity  $k$ .

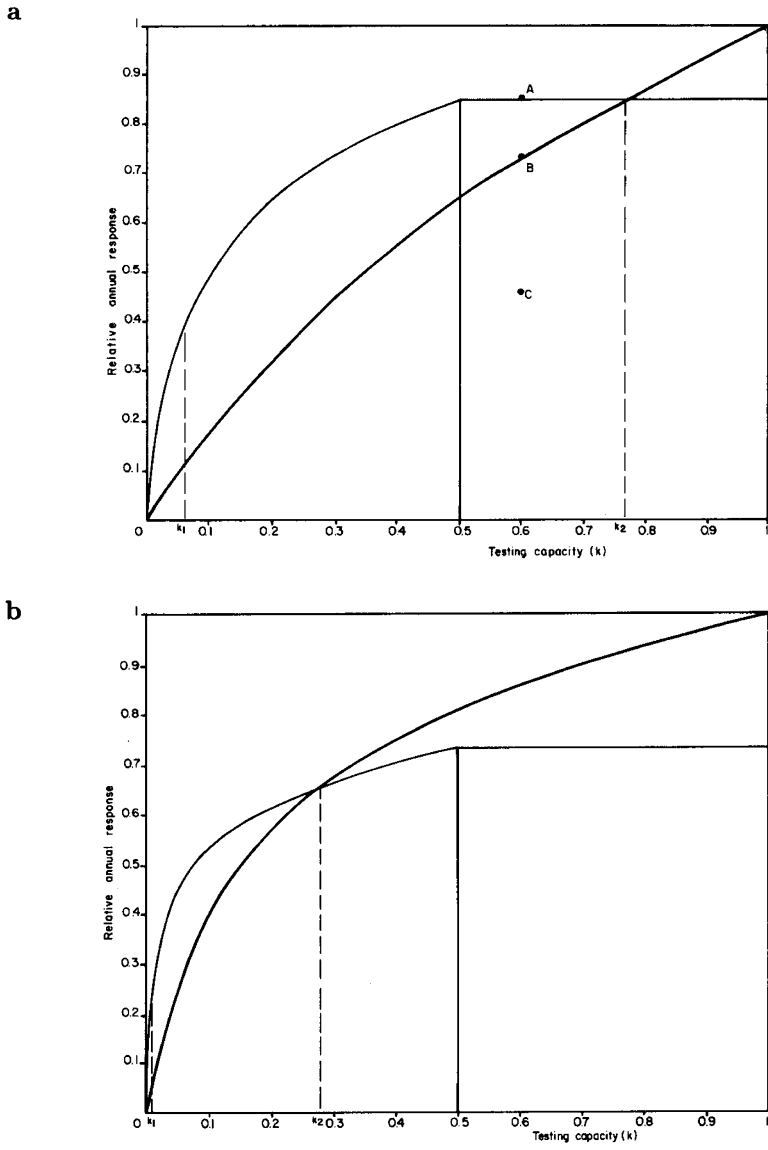
one of the patterns illustrated in Fig 1, according to whether  $k_2 < 0.5$  or  $k_2 > 0.5$ . In the latter case, rather paradoxically, the extra space available when all males are tested should not be used for testing. The worst solution would actually be to use it for testing females, as shown by point C in Fig 1a. This is because the extra selection intensity obtained by testing females is more than offset by the increase in their generation interval.

## DISCUSSION

A parallel can be drawn between the above results and those of Smith (1969). He considered maximizing selection intensity, or response per generation, for a given number of testing places ( $T$ ) available, assuming a fixed generation interval. Here the objective is to maximize annual response, with variable generation length, and the testing capacity ( $k$ ), which is defined on a yearly basis. If generation interval is set at a value  $t$ , and  $T$  is defined as the number tested per breeding female over a period of time equal to the average breeding life of sires and dams,  $2(t - a)$ , the relationship between  $T$  and  $k$  is:

$$T = 4k\lambda(t - a) \quad (11)$$

In case 1, the selection strategy recommended here, may be compared to the rule given by Smith (1969), which states that "if testing facilities are very limited, it is better to use untested males, than males which are below average". Thus,  $1/2$  is the maximum proportion to select in order to maximize the response per generation, as against  $1/3.7$  (two), if the response per year is considered. The two approaches can, for instance, be compared in terms of expected annual response for a testing capacity equal to  $k_1$ . Using Smith's approach, the critical number of testing places below which untested males should be used, is  $T = 2/c$ , i.e., 2 male candidates tested per sire to be replaced. This implies a generation interval  $t = 2.1a + 1/3.7\lambda$ , a value obtained from solving eqn(11) for  $T = 2/c$  and  $k = k_1$ , and which can also be derived from eqn(2c) with  $m = 1$ ,  $n_2 = 2$  and



**Fig 1.** Expected annual response to mass selection as a function of testing capacity ( $k$ ), expressed relative to the value for complete testing ( $k = 1$ ): — testing capacity equally allocated to the 2 sexes; - testing capacity used only for testing males.  
**(a)**  $c = 10$ ,  $a\lambda = 0.5$  Expected relative response when  $k = 0.6$  with 3 different allocations: (A) testing capacity only used for testing all male candidates available, extra space left unused; (B) testing capacity equally allocated to the 2 sexes; (C) testing capacity used for testing all males candidates available, extra space allocated to female candidates.  
**(b)**  $c = 10$ ,  $a\lambda = 5$  ( $c$ ,  $a$ ,  $\lambda$ ,  $k_1$  and  $k_2$  are as defined in Table I.)

$l_2 = 2k_1\lambda$ . The supplementary gain expected from applying Table I strategy when  $k = k_1$ , using the value  $R_a = 1.2\lambda/(1 + 4a\lambda)$  derived from eqn(7), can then be shown to range from 33 to 53%, when  $a\lambda$  goes from 0.5 to 5.

For case 2 and 3, Smith's approach leads to recommendation of a gradual increase in the proportion ( $f$ ) of females selected, whereas, here no intermediate optimum for  $f$  exists, but rather an abrupt change from  $f = 0$  to  $f = 1$ , between case 2 and case 3. In Smith's approach, the gradual increase in  $f$  should start at  $T \simeq 1$ , and case 3 is reached when  $T \simeq 3$ . Taking a generation interval  $t = 2a$ , usual in livestock populations, it can be seen from eqn(11) that the equivalent testing capacity necessary to reach case 3 is  $k = 3/4a\lambda$ . This means that when the generation interval is not acted upon, case 3 can be reached only if  $a\lambda > 0.75$ .

The model used in this study, rests on several simplifying assumptions, of which a detailed discussion has been given by Ollivier (1974). The population undergoing selection is supposed to be large and stationary in size, and a uniform age distribution of the breeding animals is assumed. It perhaps should be stressed that testing space is defined relative to a given population size. If the testing space were defined as an absolute value, the population size might be reduced to match the testing capacity in order to increase the immediate response. Loss of genetic variance would, however, be incurred, thus compromising long-term response. Optimal strategies for maximizing long-term response to selection in such situations have been explored by Robertson (1960; 1970), Smith (1969; 1981), and James (1972), among others, under the assumption of a fixed generation time. In a situation of restricted yearly testing facilities, it would then be advisable to maximize the generation length in order to also maximize the number of candidates per generation.

The assumption of a uniform age distribution is not generally met in practice and can only be accepted as an approximation in situations of fast replacement, or when the increase in fecundity with age can compensate for the gradual decay in the number of breeding individuals. With low testing capacity, however, the procedure recommended here, implies contrasting turnover rates between males (selected) and females (taken at random). When  $k < k_2$ , the female generation interval should be minimized, whereas, the male generation interval will gradually increase as  $k$  decreases. When  $k = k_1$ , this interval exceeds 3 times the age at birth of the first offspring, as shown in Table I. Obviously, such a strategy can be strictly implemented, provided enough semen from the selected sires can be stored and if breeding can be carried on, artificially if necessary.

In spite of the limitations discussed above, the results presented may serve as guidelines for the optimal use of limited testing facilities. They also show that a sizeable proportion of the maximum genetic gain can be obtained with very limited testing facilities. The conclusions are restricted to mass selection, which requires no pedigree information. Extension to family selection may be considered. Extra selection intensity could, for instance, be obtained by applying family selection to untested candidates whenever tested relatives are available. However, in a situation of limited testing facilities, information on relatives would also be limited and the extra selection intensity would have to be set against the resulting increase in generation interval. Evaluation of tested candidates could also be made more accurate by using combined selection, as in the designs considered by Poujardieu



and Rouvier (1971) or, more generally, with best linear unbiased prediction of breeding value (Henderson, 1963). One would then expect the accuracy of selection to increase with the testing capacity. This would mean, for any given testing capacity, a lower response relative to complete testing, than with mass selection. An optimal strategy would, therefore, be more complex to establish for combined selection, as it would depend on the relationship between the testing capacity and the selection accuracy. x??

## REFERENCES

- Dickerson GE, Hazel LN (1944) Effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. *J Agric Res* 69, 459-476
- Henderson CR (1963) Selection index and expected genetic advance. In: *Statistical Genetics and Plant Breeding* (Hanson WD, Robinson HF, eds) NASNRC Publ. 982, 141-163
- James JW (1972) Optimum selection intensity in breeding programmes. *Anim Prod* 14, 1-9
- Ollivier L (1974) Optimum replacement rates in animal breeding. *Anim Prod* 19, 257-271
- Ollivier L (1988) Current principles and future prospects in selection of farm animals. In: *Proceedings of the Second International Conference on Quantitative Genetics* (Weir BS, Eisen EJ, Goodman MM, Namkoong G, eds) Sinauer Associates, Sunderland, MA, 438-450
- Poujardieu B, Rouvier R (1971) Optimisation du plan d'accouplement dans la sélection combinée. *Ann Génét Sél Anim* 3, 509-519
- Robertson A (1960) A theory of limits in artificial selection. *Proc R Soc, Ser B*, 153, 234-249
- Robertson A (1970) Some optimum problems in individual selection. *Theor Pop Biol*, 1, 120-127
- Smith C (1969) Optimum selection procedures in animal breeding. *Anim Prod* 11, 433-442
- Smith C (1981) Levels of investment in testing and genetic improvement of livestock. *Livest Prod Sci* 8 193-201