

## Single and two-stage selection on different indices in open nucleus breeding systems

J.P. MUELLER (\*)

*School of Wool and Pastoral Sciences, The University of New South Wales  
Kensington, N.S.W., Australia, 2033*

### Summary

When different information is available from the various parts of a population the construction of correspondingly different selection indices is required. Selection criteria used for males may differ from those for females and in open nucleus systems, indices for the nucleus may not apply for the base. In order to test the effects of alternative allocation of selection efforts and to find the optimum breeding design in each case, formulae were adapted to predict the rate of genetic gain in open nucleus systems with varying selection criteria. Selection on different index sets may occur in one or two stages including progeny testing. Evaluations for a range of arbitrarily chosen index sets indicates that the genetic gain in a nucleus system is particularly sensitive to changes in the relative accuracy of indices used for sires and for base females. A similar improvement of selection accuracy of sires and base females increases genetic gain by 20-45 p. 100 and 10-20 p. 100 respectively. The higher limit is achieved when selection accuracy in the opposite sex is low. In two-stage programs and progeny testing schemes, results depend on the relative accuracy of indices used in the two stages. When sires are most accurately evaluated, opening the nucleus adds little to the gain in the system, whereas accurate selection of base born females in a single stage or after a first screening makes the open nucleus structure very attractive. The results are used to compare alternatives and optimise design in a simple sheep example.

*Key words* : Population structure, index selection, progeny test, open nucleus system.

### Résumé

*Sélection en une et deux étapes sur différents indices en systèmes  
à noyau ouvert*

L'existence d'informations spécifiques à des sous-ensembles d'une même population implique la construction d'indices de sélection différents. Les critères de sélection appliqués aux mâles peuvent différer de ceux relatifs aux femelles ; de même, dans les systèmes à noyau ouvert, les indices définis pour le noyau peuvent être inapplicables à la population

---

(\*) Present address : Instituto Nacional de Tecnología Agropecuaria, Casilla de Correo 277, San Carlos de Bariloche 8400, Argentina.

de base. Aussi, en vue d'étudier l'incidence de différentes politiques de sélection et d'optimiser celles-ci, des formules ont été mises au point qui expriment le gain génétique en système à noyau ouvert, en fonction du critère de sélection utilisé. Une sélection sur différents jeux d'indices peut survenir en une ou deux étapes y compris celle du contrôle de descendance. L'étude d'une gamme d'indices de sélection arbitrairement choisis montre que le gain génétique en système à noyau est particulièrement sensible à des variations de la précision relative des indices appliqués aux mâles d'une part, et aux femelles de la base d'autre part. Une amélioration équivalente de la précision de sélection des mâles et des femelles conduit à un accroissement du gain génétique d'environ 20 à 45 p. 100 et 10 à 20 p. 100 respectivement. Un plafond est atteint quand la précision de sélection dans le sexe opposé reste à un niveau faible.

Pour les programmes de sélection en 2 étapes basés sur le contrôle de la descendance, les résultats dépendent de la précision relative des indices appliquées à chacune des étapes. Si les pères sont connus précisément, l'ouverture du noyau n'entraîne qu'un faible gain génétique ; au contraire, l'application d'une sélection précise des femelles, nées dans la base, qu'elle soit appliquée en une seule étape ou après un tri initial, rend la structure en noyau ouvert très attractive. Ces résultats sont appliqués à la comparaison et à l'optimisation de programmes de sélection de l'espèce ovine.

*Mots clés : Structure de population, sélection sur indice, contrôle de descendance, système à noyau ouvert.*

## I. Introduction

Once the breeding objective has been defined, a breeder has to choose suitable selection criteria and design the breeding program. Maximum response to selection is obtained if all available information is used in a selection index. Different indices must often be constructed because the information available may vary among different parts of a structured population. This is of particular interest in the evaluation of hierarchical systems with upward gene migration (open nucleus systems) since, in these, selection in the lower levels contributes to genetic gain of the whole system.

JAMES (1977) developed formulae to predict genetic gains in open nucleus systems and evaluated such systems assuming the same selection criterion was used in both layers and sexes. HOPKINS (1978) showed that adopting strategies which concentrate selection efficiency in the nucleus may increase the rate of genetic response if the system is designed appropriately. Thus, in terms of index selection we may have different test accuracies in the nucleus and base : such situations have also been discussed in the context of British cattle group breeding schemes by GUY & STEANE (1980).

Another possibility is that selection criteria could vary between sexes in the nucleus, and between nucleus and base females. Indeed, a further point worth considering is two-stage selection in base females. In open nucleus systems large numbers of base females must be measured, of which a very small proportion will be used as nucleus replacements. In the large sheep flocks of the Southern Hemisphere this is at once the key advantage and the major problem of an open nucleus system, since the cost of measurement prohibits the collection of detailed information on all base females. If preliminary selection could be made on measurements cheap to obtain, followed by a second selection on more expensive criteria obtained for only a small fraction of base females, the extra genetic gain might compensate for the additional costs. Similarly, sires could be selected in stages, since it may be impracticable to retain all of them until full information is collected. Two-stage selection has not yet

been evaluated in the context of open nucleus systems, though it seems to be a promising alternative to take into account. A special case of two-stage selection of sires arises when the second stage includes progeny test results.

We may generalise evaluations of open nucleus systems for single and two-stage selection by first rewriting the basic equation in a form helpful for consideration of selection using different indices. Predicted response to selection in nucleus systems with more than one index can then be used to define the optimum breeding design. The sensitivity of genetic gain in such systems to changes in the accuracy of selection of different sections of the population may indicate a rational distribution of effort in collecting data for the construction of different indices.

The aim of this study is to provide explicit methods for evaluating selection in open nucleus systems with varying selection criteria, rather than exploring particular situations. Examples are given to illustrate application of the methods, not for their intrinsic interest. The complexity of such systems requires that a large number of symbols are used to describe them. These symbols are defined in the text and summarised in an Appendix.

## II. Methods

### A. Selection based on a single index

Suppose the aim is to improve aggregate genotype  $G$  by selection on an index  $I$ . In an open nucleus system  $G$  will be the same for base and nucleus. It is well known that the best index is given by the multiple regression of  $G$  on the traits in the index, and that the genetic superiority of a selected group is  $s(q)r_{GI}\sigma_G$  where  $s(q)$  is the standardised selection differential achieved by selecting the best fraction  $q$  of a normal distribution,  $r_{GI}$  is the correlation of index and breeding objective and  $\sigma_G$  is the standard deviation of  $G$  values.

Correlations are unaffected by scale changes so that choosing the regression of  $G$  on  $I$  as unity the genetic gain in breeding value is  $s(q)\sigma_I$  where  $\sigma_I$  is the standard deviation of index values. In what follows response to selection is calculated in the latter form, one unit change in the index corresponding to a unit change in breeding value. In practice other scales may be used, but  $\sigma_I$  must then be interpreted as  $r_{GI}\sigma_G$ .

JAMES (1977) gave a general expression for the steady state genetic gain per generation in an open nucleus system in which all sires are selected from nucleus-born males, a fraction  $x$  of nucleus dams are born in the base, and a fraction  $y$  of base dams are born in the nucleus. The total proportions of males and females selected are denoted  $a$  and  $b$  respectively, and generation length is assumed equal in nucleus and base.

$$(1) \Delta G = \frac{(1+y)(i_{MN} + (1-x)i_{NFN} + xi_{BFN}) + x(i_{MB} + (1-y)i_{BFB} + yi_{NFB})}{2(1+y+x)} \sigma_I$$

Standardised selection differentials for males used in the nucleus ( $i_{MN}$ ) and the base ( $i_{MB}$ ) are :

$$(2) \quad i_{MN} = s(a); \quad i_{MB} = (s(a/p) - pi_{MN}) / (1 - p)$$

where  $p$  is the proportion of the population in the nucleus. The remaining selection differentials are for females. For example,  $i_{BFN}$  is the differential for base-born females used in the nucleus, and so on.

Appropriate selection differentials can be obtained by noting that the proportions to be selected in each case are :

$$\begin{aligned} q_{NFN} &= b(1 - x); & q_{NFB} &= by(1 - p) / p; \\ q_{BFN} &= bxp / (1 - p); & q_{BFB} &= b(1 - y). \end{aligned}$$

Writing the total proportions selected in the nucleus and in the base as  $q_{NFT}$  and  $q_{BFT}$ , then :

$$(3) \quad i_{NFN} = s(q_{NFN}); \quad i_{NFB} = (q_{NFT} s(q_{NFT}) - q_{NFN} i_{NFN}) / q_{NFB}$$

$$(4) \quad i_{BFN} = s(q_{BFN}); \quad i_{BFB} = (q_{BFT} s(q_{BFT}) - q_{BFN} i_{BFN}) / q_{BFB}$$

### B. Different indices in the same system

In this section we consider the case where base females are selected on index  $I_B$ , nucleus females on index  $I_N$  and males are selected on index  $I_M$ . Multiplying the selection differentials in equation (1) by the standard deviations of the corresponding indices ( $\sigma_B$ ,  $\sigma_N$  and  $\sigma_M$  respectively) and collecting terms we find :

$$(5) \quad \Delta G = w_M \sigma_M + w_N \sigma_N + w_B \sigma_B$$

with weights :

$$\begin{aligned} w_M &= ((1 + y) i_{MN} + xi_{MB}) / g \\ w_N &= ((1 + y)(1 - x) i_{NFN} + xy i_{NFB}) / g \\ w_B &= x((1 + y) i_{BFN} + (1 - y) i_{BFB}) / g \end{aligned}$$

where  $g$  is  $2(1 + y + x)$ . It is worth noting that in a closed nucleus system  $x = 0$  and equation (5) reduces to  $(i_{MN} \sigma_M + i_{NFN} \sigma_N) / 2$ .

### C. Two-stage selection of base females

Suppose selection in the base occurs in stages. In the first stage a proportion  $q_1$  is selected by truncating the standardised distribution of index values  $I_{B1}$  at point  $t$ . The change in breeding value after this first stage is  $s(q_1) \sigma_{B1}$ . As will be seen later the fraction selected in the first stage is normally less than the total proportion of females required as replacements ( $q_{BFT}$ ). Hence the next best proportion  $q_{BFT} - q_1$  is used as replacement in the base. The remainder  $1 - q_{BFT}$  is culled. Among those individuals accepted for the second stage a fraction  $q_2$  is selected on the more accurate index  $I_{B2}$ . It is assumed that information from the first stage is used in the second. Since a fixed proportion of base females is required for the nucleus ( $q_{BFN}$ ), the proportions  $q_1$  and  $q_2$  are not independent ( $q_1 q_2 = q_{BFN}$ ).

A good approximation of the gain from the second stage selection, due to COCHRAN (1951), assumes  $I_{B2}$  and  $G$  remain jointly normally distributed after the first selection. Writing  $i_1 = s(q_1)$  and  $i_2 = s(q_2)$  the genetic differential of change in mean breeding value in the fraction of base females for the nucleus is :

$$D_{BFN} = i_1\sigma_{B1} + i_2\sigma_{B2} \sqrt{1 - r^2c},$$

where  $r$  is the correlation between indices used in the two stages. Since  $I_{B1}$  and  $I_{B2}$  are constructed with the same breeding objective,  $r = \sigma_{B1}/\sigma_{B2}$ . The factor  $1 - r^2c$  is the proportion of the variance in  $I_{B2}$  left in the group saved for further measurements, thus  $c = i_1(i_1 - t)$ . Making use of tables of the bivariate normal distribution it can be shown that the approximation holds well unless  $r$  is close to unity in which case the second stage would become worthless.

To find the genetic differential for base dam replacements, we recall that all surplus animals from the second stage selection are used in the base :

$$D_{BFB} = (q_{BFT} s(q_{BFT}) \sigma_{B1} - q_{BFN} D_{BFN}) / q_{BFB}$$

The genetic gain in a nucleus system with two-stage selection of base females can be calculated by replacing  $i_{BFN} \sigma_B$  and  $i_{BFB} \sigma_B$  in equation (5) with  $D_{BFN}$  and  $D_{BFB}$  from above such that :

$$(6) \quad \Delta G = w_M \sigma_M + w_N \sigma_N + w_{B1} \sigma_{B1} + w_{B2} \sigma_{B2} \sqrt{1 - r^2c},$$

where :

$$w_{B1} = x(i_1(1 + y - (1 - y) q_{BFN}/q_{BFB}) + (1 - y) s(q_{BFT}) q_{BFT}/q_{BFB})/g.$$

$$w_{B2} = xi_2(1 + y - (1 - y) q_{BFN}/q_{BFB})/g$$

#### D. Two-stage selection of sires

Similarly to the previous case, consider sire selection in two stages. First a proportion  $q_1$  (now we use  $q_1$  and  $q_2$  for the selection of males) is selected on index  $I_{M1}$ . Among these a proportion  $q_2$  is selected on the second index  $I_{M2}$ . The restriction is  $q_1q_2 = a$ , the final proportion required for the nucleus.

The genetic selection differential of males for the nucleus is :

$$D_{MN} = i_1\sigma_{M1} + i_2\sigma_{M2} \sqrt{1 - r^2c}.$$

The term in the square root has its equivalent in equation (6). If  $q_1$  is less than the total proportion of males required ( $a/p$ ) we find the genetic selection differential of males for the base as :

$$D_{MB} = (s(a/p) \sigma_{M1} - pD_{MN}) / (1 - p).$$

Suppose  $q_1 \geq a/p$  which might be taken when, for instance, artificial insemination is used and, therefore, only very few sires are needed. In this case :

$$D_{MB} = i_1\sigma_{M1} + (s(q_2/p) - pi_2) \sigma_{M2} \sqrt{1 - r^2c} / (1 - p).$$

We may write the equation for the steady state genetic gain with two-stage selection of sires in the convenient form :

$$(7) \quad \Delta G = w_N \sigma_N + w_B \sigma_B + w_{M1} \sigma_{M1} + w_{M2} \sigma_{M2} \sqrt{1 - r^2c}.$$

The weights  $w_{M1}$  and  $w_{M2}$  are derived by simply substituting  $D_{MN}$  and the proper  $D_{MB}$  for  $i_{MN}\sigma_M$  and  $i_{MB}\sigma_M$  in equation (5).

### E. Progeny testing

We regard progeny testing as implying a nucleus system in that the female population is divided in two groups, the nucleus in which all prospective sires are born and the base in which a proportion  $q_1$  selected on index  $I_{M1}$ , of these young sires is tested. In a second stage, a fraction  $q_2$  is chosen on the combined information of individual and progeny performance (index  $I_{M2}$ ) and mated in the nucleus. The traits need not be the same and may be sex-limited.

For a given fertility level  $f$  in the system and mating ratios in the base ( $M_B$ ) and in the nucleus ( $M_N$ ) expressed as sires/dams, we have from MUELLER & JAMES (1983) the proportions :

$$q_1 = 2(1-p)M_B/pf, \quad q_2 = pM_N/M_B(1-p).$$

It is assumed that sires are used an equal number of times in nucleus and base.

The genetic differential for young sires is :

$$D_{MB} = i_1\sigma_{M1}$$

and for sires accepted after the second stage it is :

$$D_{MN} = D_{MB} + i_2\sigma_{M2} \sqrt{1-r^2c}.$$

The genetic gain in the steady state situation of the system is described by equation (7) with  $w_{M1} = i_1/2$  and  $w_{M2} = i_2(1+y)/g$ .

### F. Evaluation of formulae

The total proportion of males and females required as replacements (a and b) are usually characteristic of a particular population and to a large extent uncontrollable except for the use of artificial insemination and changes in age structure. The breeder can, however, manipulate the structure of the breeding population by choosing the size of the nucleus ( $p$ ) and by deciding on the proportion of individuals transferred between base and nucleus ( $x$  and  $y$ ). Since we expect selection in the nucleus to be at least as accurate as in the base the optimum nucleus size is small (JAMES, 1977) and  $y$  is then necessarily small. Thus with little or no loss of efficiency we assume that all surplus females from the nucleus are used in the base and restrict our attention to the more relevant design parameters  $p$  and  $x$ .

In order to quantify the response to selection for several combinations of indices we use equations (5), (6) and (7) for given a and b over a range of  $x$  and  $p$ . In the case of progeny testing, annual response rate is calculated in a population which requires 70 p. 100 of females for replacements ( $b = 0.7$ ), with fertility level at 80 p. 100 ( $f = 0.8$ ), mating ratio in the nucleus of 0.4 p. 100 ( $M_N = 0.004$ ), and mating ratio in the base of 2 p. 100 ( $M_B = 0.02$ ). The test is based on  $f/M_B = 40$  offspring of both sexes per young sire. Age structure of females is the same in nucleus and base. Young sires are used once in the base and those selected on the progeny test are

used once in the nucleus. Age at first offspring is 2 years in both sexes. It is not implied that age structures and mating ratios considered in the evaluations are always optimal. The situation described could apply to a sheep population in which the base is run under extensive conditions and the nucleus having artificial insemination facilities.

For cases other than progeny testing results are on a per generation basis. The implications of this assumption will be discussed later.

### III. Results

#### A. Design of nucleus systems with single stage selection

A particular set of indices is named by a 3 digit number representing the standard deviations (or accuracies) of indices applied to select males, females in the nucleus and females in the base. Index set 333 serves as reference and, for example, index set 231 is a case where  $\sigma_M = 2$ ,  $\sigma_N = 3$ , and  $\sigma_B = 1$ .

Before analysing particular index sets, we might see how the three weights  $w_M$ ,  $w_N$  and  $w_B$  in equation (5) change with increasing proportion of base born females in the nucleus (fig. 1).

The weight for males is large, especially when intense use of them is possible. Thus the open nucleus structure would become unfavourable when artificial insemination is extensively used and the ratio  $\sigma_M/\sigma_B$  is large.

The optimum size of the nucleus for several index sets is shown in figure 2. The population considered with replacement proportions  $a = 0.05$  and  $b = 0.7$  is typical of many sheep and cattle populations. With  $\sigma_M > \sigma_B$  as in 321 a large nucleus would be suggested whereas the opposite is appropriate with 133.

A comparison of maximum response rates in the figures shows the relative effects of the application of different indices. It is clear that optimisation of nucleus size is of little effect compared to changes in the index sets.

The determination of the number of females which should be transferred to the nucleus is of greater relevance (fig. 3).

With large  $\sigma_M$  and small  $\sigma_B$ ,  $x$  should be small. The open nucleus structure with large upward transfers becomes particularly efficient with index sets like 233 and 133. The control set (333) yields optima as in figures 2 and 3 of JAMES (1977).

#### B. Design of nucleus systems with two-stage selection on individual performance

For both, the two-stage selection of base females and the two-stage selection of sires, we need first to consider the allocation of selection intensity to each stage, that is, we need to find values of  $q_1$  and  $q_2$  so as to achieve near maximum

gains. Maximum efficiency would, of course, be to select the whole population on the full index ( $I_2$ ) taking  $q_1 = 1$ . However, we want to measure only a small proportion on  $I_2$  and still achieve a high fraction, say 80 or 95 p. 100 of the possible gain in  $D_{BFN}$ .

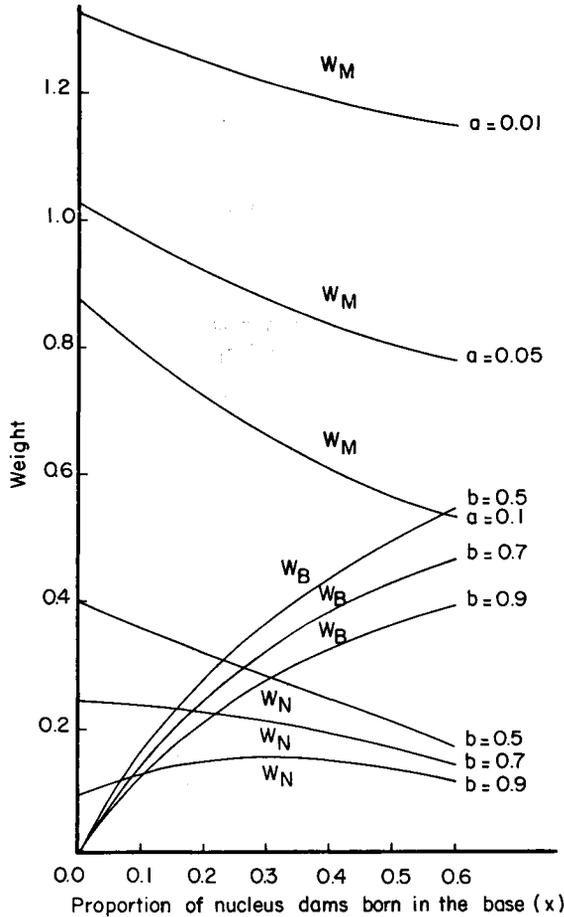


FIG. 1

*Effect of the proportion of nucleus dams born in the base ( $x$ ) of a nucleus system, on the weighted selection differentials ( $w$ ) achieved for males ( $M$ ), base born females ( $B$ ) and nucleus born females ( $N$ ) in populations with different overall fractions of males ( $a$ ) and females ( $b$ ) required for breeding.*

*Incidence de la proportion ( $x$ ) de mères du noyau nées dans la base sur les différentielles de sélection pondérées obtenues pour les mâles ( $M$ ), les femelles nées dans la base ( $B$ ), celles nées dans le noyau ( $N$ ) dans des populations à fractions variables de mâles ( $a$ ) et de femelles ( $b$ ) nécessaires à la reproduction.*

Cochran suggested that selecting equal proportions in a multi-stage program is approximately optimum. Numerical evaluation shows that for a range of final proportions required ( $q_1q_2$ ) a nearly constant efficiency is achieved in this way, but the magnitude of the efficiency depends on the correlation  $r$  between indices used (fig. 4).

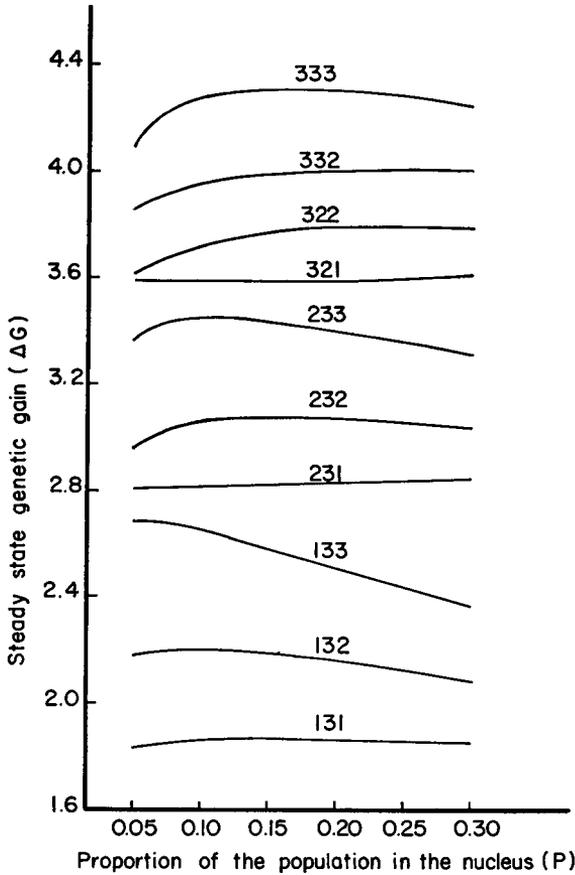


FIG. 2

Effect of nucleus size ( $p$ ) on the steady state genetic gain of a nucleus system in which sires, nucleus born dams and base born dams are selected on indices with relative accuracies given by the digits in that order (example for  $a = 0.05$ ,  $b = 0.7$  at optimum  $x$ ).

*Effet de la taille relative ( $p$ ) du noyau sur le gain génétique asymptotique dans un système à noyau où les pères, les mères, nées dans le noyau et celles nées dans la base sont sélectionnés sur des indices de précision relative donnée par les chiffres présentés dans cet ordre (exemple pour  $a = 0,05$ ,  $b = 0,7$  à l'optimum  $x$ ).*

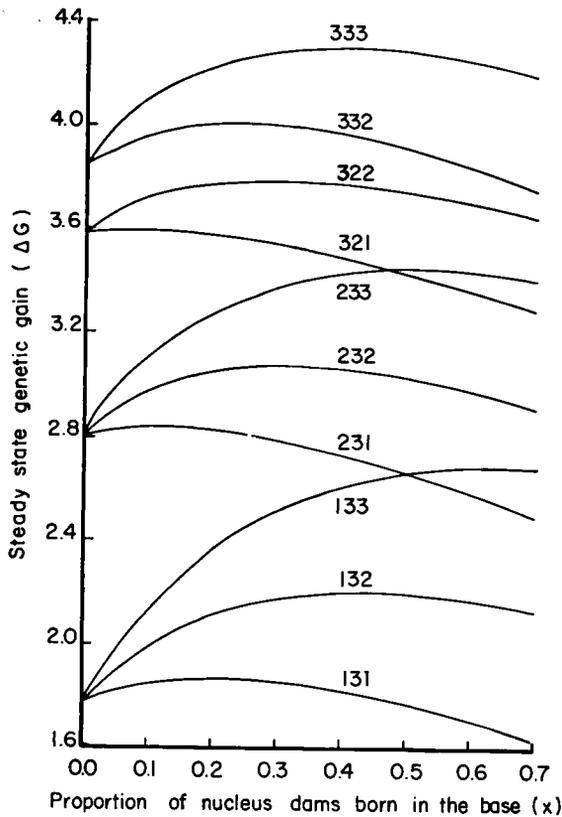


FIG. 3

*Effect of the proportion of nucleus dams born in the base ( $x$ ) on the steady state rate of genetic gain of a nucleus system in which sires, nucleus dams and base dams are selected on indices with relative accuracies given by the digits in that order (example for  $a = 0.05$ ,  $b = 0.7$  at optimum  $p$ ).*

*Effet de la proportion de mères du noyau ( $x$ ) qui sont nées dans la base sur le gain génétique asymptotique attendu dans un système à noyau où les pères, les mères du noyau et celles de la base sont sélectionnées sur des indices de précision relative donnée par les chiffres présentés dans cet ordre (exemple pour  $a = 0,05$ ,  $b = 0,7$  à l'optimum  $p$ ).*

For values of  $r$  in the range  $1/3$  to  $2/3$  selecting equal proportions ( $q_1 = q_2$ ) achieves about 80 p. 100 of the possible gain. This simple criterion will be used in the examples of two-stage selection on individual performance.

Results for two-stage selection of base females are shown in figure 5 and results of two-stage selection of males in table 1. The standard deviations of indices used for

the first and second stage appear in brackets at the beginning of a set ( $\sigma_{M1}\sigma_{M2}$ ) or at the end ( $\sigma_{B1}\sigma_{B2}$ ).

Results on optimum design are as expected. As selection accuracy increases with the second stage the design parameter shift accordingly towards larger nucleus and smaller base contribution when the second index is applied to males and vice versa when base born females are selected in two stages.

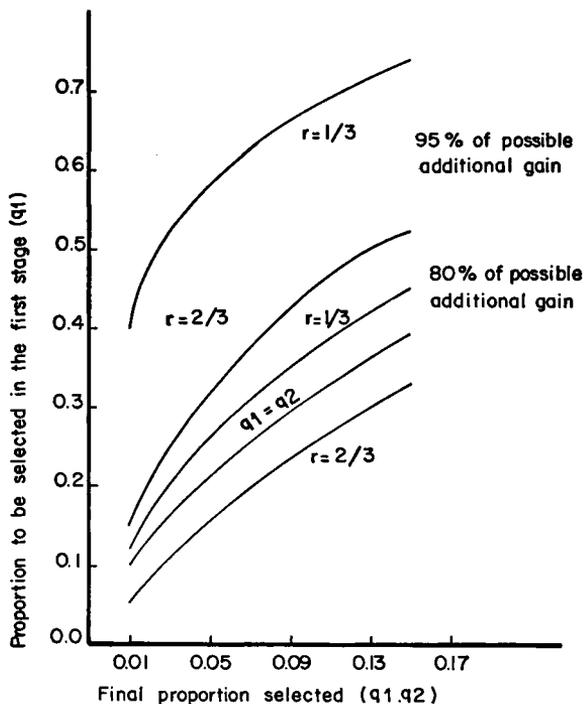


FIG. 4

Proportions of a population to be selected in the first stage ( $q_1$ ) of a two-stage program which achieve 80 p. 100 or 95 p. 100 of the possible additional gain due to a second stage selection for different final proportions required ( $q_1, q_2$ ).

Equal allocation of selection ( $q_1 = q_2$ ) achieves about 80 p. 100 of possible additional gain when selection in the first stage is about half as accurate as in the second stage ( $r = \sigma_{11}/\sigma_{12} = 0.5$ ).

Taux de sélection ( $q_1$ ) à appliquer à la première étape d'un programme de sélection en 2 étapes qui procure 80 à 95 p. 100 du gain génétique potentiel supplémentaire dû à la sélection en une 2<sup>e</sup> étape pour différentes valeurs du taux final ( $q_1, q_2$ ). Une répartition égale de la pression de sélection ( $q_1 = q_2$ ) conduit à 80 p. 100 du gain marginal potentiel si la sélection au 1<sup>er</sup> stade est moitié moins précise que celle appliquée au 2<sup>e</sup> stade ( $r = \sigma_{11}/\sigma_{12} = 0,5$ ).

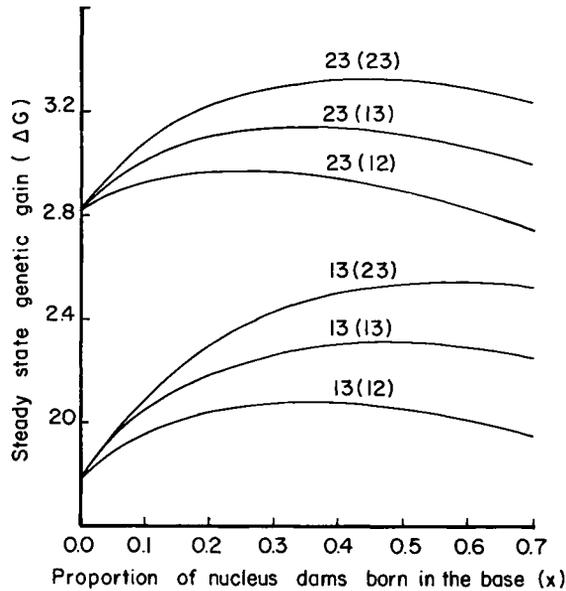


FIG. 5

Effect of the proportion of nucleus dams born in the base ( $x$ ) on the steady state genetic gain of a nucleus system in which sires, nucleus born dams and base born dams are selected on indices with relative accuracies given by digits in that order. The fourth digit corresponds to the accuracy of the index used in a second-stage selection of base born females (example for  $a = 0.05$ ,  $b = 0.7$ ,  $q_1 = q_2$  at optimum  $p$ ).

Effet de la proportion de mères du noyau ( $x$ ) qui sont nées dans la base sur le gain génétique asymptotique attendu, dans un système à noyau où les pères, les mères nées dans le noyau et celles nées dans la base sont sélectionnées sur des indices de précision relative donnée par les chiffres présentés dans cet ordre.

Le 4<sup>e</sup> chiffre, correspond à la précision de l'indice appliqué en 2<sup>e</sup> étape sur les femelles nées dans la base (exemple pour  $a = 0,05$ ,  $b = 0,7$ ,  $q_1 = q_2$  à l'optimum  $p$ ).

TABLE 1

Ratio of genetic gain from two-stage selection of males on indices  $I_{M1}$  and  $I_{M2}$  to that from single-stage selection on  $I_{M1}$ .

Rapport des gains génétiques attendus par une sélection en 2 étapes des mâles sur des indices  $I_{M1}$  et  $I_{M2}$  d'une part, et par celle en une seule étape sur  $I_{M1}$  d'autre part.

		$(\sigma_{M1}\sigma_{M2})$		
		(13)	(12)	(23)
$\sigma_N\sigma_B$ .....	3 1	1.78	1.41	1.30
$\sigma_N(\sigma_{B1}\sigma_{B2})$ .....	3 (13)	1.51	1.26	1.23
$\sigma_N\sigma_B$ .....	3 3	1.38	1.20	1.18

In the nucleus, females are selected on  $I_N$ , and in the base on  $I_B$  in a single stage or on  $I_{B1}$  and  $I_{B2}$  in two stages.  $\sigma$  denotes the standard deviation (accuracy) of the subscripted indices.

Dans le noyau, les femelles sont sélectionnées sur  $I_N$ , dans la base, soit sur  $I_B$  en une seule étape, soit sur  $I_{B1}$  et  $I_{B2}$  en 2 étapes.  $\sigma$  désigne l'écart-type (précision) des indices.

C. Design with progeny testing

In this case the proportion selected in the first stage ( $q_1$ ) is a function of the nucleus size for a given mating ratio in the base. The number of progeny on which the test is based is independent of  $p$ . Therefore the nucleus should be very small to allow a high selection intensity on  $I_{M2}$  when this index is relatively accurate. Population size will set a minimum on  $p$  compatible with inbreeding considerations, but also pedigree recording may increase beyond possibilities as the base becomes larger. Consequently a restricted range of  $p$  needs to be taken in the evaluations. Table 2 shows the efficiency of progeny testing (as the ratio of  $\Delta G$  at optimum  $p$  and  $x$  for progeny test and individual selection) when at least 10  $p$ . 100 of the population is mated to tested sires ( $p = 0.1$ ).

TABLE 2

Ratio of genetic gain from progeny testing with accuracy  $\sigma_{M2}$  of individually selected males to that from individual performance selection only with accuracy  $\sigma_{M1}$ .

Rapport des gains génétiques dus à la sélection de mâles sur descendance avec une précision  $\sigma_{M2}$  d'une part, et à la seule sélection individuelle de précision  $\sigma_{M1}$  d'autre part.

Age of sires in the nucleus	$(\sigma_{M1}\sigma_{M2})$					
	(14)	(13)	(24)	(12)	(23)	(34)
3 .....	2.08	1.65	1.43	1.25	1.13	1.06
4 .....	1.80	1.45	1.24	1.11	1.00	0.93
5 .....	1.60	1.29	1.10	1.00	0.90	0.83

All females are selected on an index with accuracy equal to 2. See text for assumptions on age structure.

Toutes les femelles sont choisies sur un indice de précision égale à 2. Se reporter au texte pour connaître les hypothèses sur la structure d'âge.

TABLE 3

Approximate optimum proportion of nucleus females born in the base ( $x$ ) and mated to proven sires for progeny testing systems with different selection accuracy of young sires ( $\sigma_{M1}$ ) and proven sires ( $\sigma_{M2}$ ).

Proportion optimum approchée de femelles du noyau qui sont nées dans la base ( $x$ ) et accouplées à des pères favorablement testés dans des systèmes de contrôle sur descendance présentant une précision de sélection différente pour les jeunes mâles ( $\sigma_{M1}$ ) et ceux d'élite ( $\sigma_{M2}$ ).

Age of sires in the nucleus	$(\sigma_{M1}\sigma_{M2})$					
	(14)	(13)	(24)	(12)	(23)	(34)
3 .....	0	0.1	0.05	0.35	0.2	0.2
4 .....	0.05	0.2	0.1	0.4	0.35	0.35
5 .....	0.15	0.35	0.2	0.55	0.45	0.4

All females are selected on an index with accuracy equal 2. See text for assumptions on age structure.

Toutes les femelles sont choisies sur un indice de précision égal à 2. Se reporter au texte pour connaître les hypothèses sur la structure d'âge.

Apart from the obvious effect of increasing age of nucleus sires, it is also clear that the smaller the ratio  $\sigma_{M1}/\sigma_{M2}$  the more efficient is the progeny test. For index sets with  $\sigma_{M1} = 1$  and for  $\sigma_{M1}\sigma_{M2} = (24)$ , optimum nucleus size is the minimum considered ( $p = 0.1$ ), whereas for (23)  $p = 0.2$  and for (34)  $p$  should be around 0.3. Optimum upward transfers ( $x$ ) are given in table 3, they go from 0.5 to 0.0 depending on how much more accurate the progeny test is.

#### IV. Discussion

The equations used to predict response to selection are based on the usual assumptions of additive inheritance, multinormal distributions and constant variances and covariances. The assumption of equal generation length in base and nucleus deserves some comments in the light of the findings by HOPKINS (1978) that additional response can be achieved from optimisation of the age structure in a nucleus system. There is no particular methodological difficulty in evaluations for overlapping generations. The response formulae would need to be divided by the weighted (with transfer rates) average (over sexes) generation length in nucleus and base. It requires trial and error procedures to locate the optimum combination of nucleus size, transfer rates and number of age groups in the different parts of the system. HOPKINS (1978) determined optimum age structures for several selection strategies, including selection at one or all ages and sequential culling. The results indicated that in general fewer age groups would be recommended in the nucleus although sensitivity of genetic gains to changes in age structure becomes negligible the more efficient selection becomes in the nucleus. In relation to the present work such more efficient selection strategies have effectively the same consequences as the use of more accurate indices. Therefore, following Hopkins' results, only when selection accuracy is fairly even in the system would we be concerned with optimising age structures. If this is the case we would predict slightly higher gains and smaller transfers and nucleus sizes than those calculated assuming equal generation length in both layers.

While HOPKINS (1978) emphasised optimisation of age structures we were primarily concerned with the effects of alternative allocation of selection efforts in the different parts of an open nucleus scheme. The study shows that useful preliminary information for the design of complex breeding systems can be inferred from inspection of the weighted selection differentials ( $w$ 's) over the relevant parameters (e.g.,  $x$  and  $p$ ) and from the relative size of accuracies of the selection indices proposed for the different parts of the system. For instance, if selection accuracy of males ( $\sigma_M$ ) can be improved by one unit (from 1 to 2 or 2 to 3) an increase in the genetic gain of roughly 25 to 45 p. 100 could be achieved depending upon whether selection of base females is very accurate ( $\sigma_B = 3$ ) or not ( $\sigma_B = 1$ ). Similarly one unit improvement in  $\sigma_B$  can raise gain by up to 20 p. 100 depending on selection accuracy of sires. If the increase in selection accuracy in the base arises from a second stage selection about 10 p. 100 extra gain can be predicted. Again for two-stage selection of sires, one unit improvement over  $\sigma_{M1}$  yields 20 to 40 p. 100 higher gains. In progeny testing schemes these extra gains are diluted by the larger generation intervals.

Application of the methods described is straight-forward. As an example consider a nucleus system in a given sheep population ( $a = 0.05$  and  $b = 0.7$ ). Take the

selection objective function and variance-covariance matrices as in PONZONI (1979), but allowing for the cost of producing a lamb by halving its price. The standard deviation of breeding values  $\sigma_G$  equals 6.2, so that those readers who prefer interpretations in  $r_{GI}$  would need to divide all index sets by 6.2. Select nucleus females on clean fleece weight, 16-month body weight and wrinkle score; select sires on the former two traits alone; we find for unscaled indices  $\sigma_M = 2.4$  and  $\sigma_N = 3.2$  from standard selection index theory. Now suppose base females are selected on greasy fleece weight ( $\sigma_B = 1.5$ ). Using equation (5) or interpolating in figures 2 and 3 we find the optimum proportion of the population in the nucleus ( $p$ ) as about 25 p. 100 and the proportion of nucleus dam replacements chosen among base dam progeny ( $x$ ) at around 18 p. 100. Introducing a second stage selection in the base such that  $I_{B2} = I_N$ , we locate optimum  $q_1 = 0.19$ ,  $p = 0.13$  and  $x = 0.36$  in figures 2, 4 and 5 or using equations (6), and recalling that  $q_1 q_2 = q_{BFN} = bxp / (1 - p)$ . Fortunately, optima are not sharp and with little loss of efficiency, these parameters can be chosen at operationally convenient levels near these results. The extra gain in the system is about 9 p. 100, which would need to be balanced against the costs of additional measurements on  $(1 - p) q_1 = 17$  p. 100 of the population.

Incorporation of cost functions in the evaluations may sometimes be helpful, but it is unlikely that generally useful results could be obtained since costs and returns would be specific to any particular nucleus system. This is particularly relevant to two-stage selection where, depending on circumstances, it may not be desirable to take  $q_1 = q_2$ .

Breeding systems including progeny testing are usually expensive and substantial rewards in genetic gain must be predicted to make them feasible. Thus, a thorough analysis of alternative designs and optimisation would be required. In many cases the range of alternative designs can be narrowed by consideration of the general pattern of results in this study. For illustration consider again the sheep example. Since PONZONI's (1979) breeding objective places a high weight on a reproductive trait (number of lambs weaned), it might be of interest to include a trait highly correlated with fertility in the index for rams. Ovulation rate of daughters is an obvious possibility. Suppose rams for the base and all ewe hoggets are selected on clean fleece weight ( $\sigma_N = \sigma_B = \sigma_{M1} = 2.0$ ) and rams for the nucleus are selected in a second stage on clean fleece weight and the average ovulation rate, measured before first joining, of 20 daughters ( $\sigma_{M2} = 3.3$ ). The age of rams in the nucleus is 4 years. From table 2 we would expect an improvement of about 8 p. 100 over single stage selection. However, with additional information in  $I_{M1}$ , such as 16-month body weight and wrinkle score,  $\sigma_{M1} = 3.2$  and  $\sigma_{M2} = 4.1$ , and single stage selection becomes more efficient (about 7 p. 100). Thus even with the very optimistic genetic correlation of ovulation rate and number of lambs weaned of 0.8 assumed (other genetic and phenotypic correlations involving ovulation rate taken as zero), there is no point in progeny testing when there is scope to improve first stage selection accuracy. The conclusions drawn from this example are, of course, restricted to populations which satisfy the particular genetic parameters assumed.

*Received December 13, 1982.*

*Accepted August 5, 1983.*

### Acknowledgements

I wish to thank D<sup>r</sup> J.W. James for many helpful suggestions. Support from the Instituto Nacional de Tecnologia Agropecuaria, I.N.T.A. (Argentina) is greatly acknowledged.

### References

- COCHRAN W.G., 1951. Improvement by means of selection. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Neyman J. (ed.), 449-470, University of California Press, Berkeley.
- GUY D.R., STEANE D.E., 1980. The optimisation of group breeding schemes. *31st Annual Meeting, European Association for Animal Production, Munich, Germany, September 1-4, 1980*.
- HOPKINS I.R., 1978. Some optimum age structures and selection methods in open nucleus breeding schemes with overlapping generations. *Anim. Prod.*, **26**, 267-276.
- JAMES J.W., 1977. Open nucleus breeding systems. *Anim. Prod.*, **24**, 287-305.
- MUELLER J.P., JAMES J.W., 1984. Design and evaluation of progeny testing in open nucleus breeding systems. *Anim. Prod.*, **38** (in press).
- PONZONI R.W., 1979. Objectives and selection criteria for Australian Merino sheep. *Proceedings of the Australian Association of Animal Breeding and Genetics*, **1**, 320-336.

### Appendix

#### *Table of symbols*

- a      proportion of males required as replacements
- b      proportion of females required as replacements
- x      proportion of nucleus female replacements born in the base (upward female transfer rate)
- y      proportion of base female replacements born in the nucleus (downward female transfer rate)
- p      nucleus size as a fraction of total population
- g      =  $2(1 + y + x)$  or two times the total transfers
- f      fertility level, number of successfully reared progeny per female joined
- q      refers to proportions
- s(q)   selection differential when an upper tail area q is selected from a normal distribution
- i      standardised selection differential, used to shorten s(q)

Subscripts for q and i :

- 1 and 2 : refer to first and second stage selection
- MN and MB : males for the nucleus and males for the base
- NFN, NFB, BFN and BFB : nucleus born females used in  
the nucleus, nucleus females used in the base,  
base females used in the nucleus and base  
females used in the base

- I selection index
- G aggregate breeding value
- $\sigma_I$  standard deviation or accuracy of index I
- $\sigma_G$  standard deviation of breeding values
- $r_{GI}$  correlation of breeding values and index
- r correlation of indices used in first and second stage equal  $\sigma_{I1}/\sigma_{I2}$
- $M_N, M_B$  mating ratios (sires/dams) in nucleus and base

Index sets :

- sets of three digits say accuracies of indices used for males, nucleus females and base females ; in general :  $\sigma_M\sigma_N\sigma_B$
- two digits in brackets indicate accuracies of indices used in the first and second stage ; in general  $(\sigma_{M1}\sigma_{M2})\sigma_N \sigma_B$  or  $\sigma_M\sigma_N (\sigma_{B1}\sigma_{B2})$

- $\Delta G$  rate of genetic gain in units of breeding value.