Theor Appl Genet (1983) 67:45–52

The expected efficiencies of some methods of selection of components for inter-genotypic mixtures

A. J. Wright
Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ, England

Received February 24, 1983; Accepted June 16, 1983
Communicated by A. Robertson

Summary. The gains in yields of mixtures expected to follow various methods of selection of components from populations of randomly constituted mixtures are formulated in terms of a statistical model and a model of first-order inter-plant competition. The types of selection investigated include that among whole mixtures and among groups or individual components on a within-mixture or unrestricted basis, or on the basis of the yields of sets of mixtures to which the group or individual is common. In all cases the sizes of mixture used for selection and for measurement of gain may differ. While evaluation of components in groups or whole mixtures allows selection for component interactions, gains are lower overall because of the reduction in variance caused by grouping. Gains due to interaction are lost if the components are pooled after selection, as in a population improvement programme. Individual selection carries some risk of negative gains, but these are reduced if assessment is made on an unrestricted rather than within-mixture basis. When second and higher order competitive interactions are absent, monoculture assessment is expected to be an efficient means of selection of components for binary and tertiary mixtures.

Key words: Competition – Intergenotypic mixture – Component selection

1 Introduction

In his study of the expected effects of selection among genotypes interacting in groups, Griffing (1967, 1968) found that selection operating on groups of one size (i.e. number of components) need not necessarily lead to positive changes in the mean of the population when evaluated in groups of a different size, but also concluded that selection based on groups of the appropriate size need not be the most efficient method. These phenomena are due not only to the existence of associate effects bearing an unknown relationship to the direct effects of genotypes, but also to the fact that direct and associate effects and any interactions among them must be regarded as specific to the size of group considered. Hence response formulae contain covariances of unknown sign and magnitude.

Further resolution of this problem requires the establishment of some relationship between the parameters defined for groups of different sizes. In the case of groups interacting genetically, as do outbreeding crop plants when intercrossed to produce synthetic populations, the Mendelian model of gene inheritance and expression provides the necessary unification. For mixtures of competing genotypes, a model of inter-plant competition which permits only pairwise interactions can fulfil a similar role, and when shown to be adequate leads to simple relationships between mixture means and variances and the reciprocal of the number of components (Wright 1982).

The present paper considers some of the different possible methods of evaluation and selection of components for mixtures on both a within and between mixture basis, and develops formulae for their expected efficiencies in terms of both statistical and competition models.

2 Models

Statistical and competition models appropriate for the description of inter-genotypic mixtures were examined in an earlier paper (Wright 1982). These will be briefly described here.

2.1 Statistical model

A general model for one component of a mixture of equal proportions of m components, including interactions up to the second order, is

$$x_{i/j\ldots m} = u + v_i + \sum_{j \neq i} a_j + \sum_{j \neq i} (v\,a)_{ij} + \sum_{j \neq i} \sum_{k > j} (a\,a)_{jk} + \sum_{j \neq i} \sum_{k > j} (v\,a\,a)_{ijk} \quad (1)$$

where summation is over the $1 \ldots m$ components in the mixture. The effects are defined with respect to a large reference population of components, and are u, the mean of all components in all mixtures, v_i, the direct effect of the ith component, a_j, the associate effect of the jth, $(v\,a)_{ij}$, the interaction of these direct and associate effects, $(a\,a)_{jk}$, the average interaction of the jth and kth associates on all other components, and $(v\,a\,a)_{ijk}$, their specific effect on the ith. Error terms would also be required to describe actual component yields. There is clearly no correspondence between similar parameters defined for different sized mixtures, if only because an increase in m leads to the use of a larger number of parameters to describe a smaller area of land (i.e. 1/m part of a plot). A complete notation would therefore include m as a subscript to all parameters to emphasise this distinction.

2.2 Competition model

This model assumes that there are no pairwise or higher-order interactions of plants with respect to the competitive influences they exert on others. Thus

$$x_{i/j\ldots m} = \sum_{j}^{m} c_{i/j}/m^2, \quad \text{and} \quad y_{ij\ldots m} = \sum_{i}^{m} \sum_{j}^{m} c_{i/j}/m^2$$

where $y_{ij\ldots m}$ is the total mixture yield and $x_{i/j\ldots m}$ that of the ith of its components as before. The parameter $c_{i/j}$ is the whole plot yield of component i in purely inter-component competition with component j (i.e. if all plants of type i were completely surrounded by plants of type j). The yield of a mixture is therefore the mean of a two dimensional array of these effects, each row of which represents the contribution of one component and includes one element which is an intra-component yield.

The equivalence of the effects defined for the competition model and those for the statistical model (1) can be established as follows. If $c_{.}$ is the mean of the intra-component effects $c_{i/i}$, over the population, $c_{i/.}$ and $c_{./i}$ are the mean direct and associate intercomponent effects for the ith variety, and $c_{./.}$ the grand mean of all inter-component effects, and $m_i = m - i$

$$\begin{aligned}
u_{(m)} &= (c_{.} + m_1\, c_{./.})/m^2, \\
v_{(m)i} &= ((c_{i/i} - c_{.}) + m_1\,(c_{i/.} - c_{./.}))/m^2, \\
a_{(m)i} &= (c_{./i} - c_{./.})/m^2, \\
(v\,a)_{(m)ij} &= (c_{i/j} - c_{i/.} - c_{./j} + c_{./.})/m^2.
\end{aligned}$$

This model, when appropriate, allows an explicit formulation of all statistical parameters for mixtures of any size. The absence of any higher order statistical effects allows the adequacy of the model to be tested using analysis of variance of the total or component yields of sets of mixtures of three or more components, and the linear relationship between mixture mean and the reciprocal of the number of components provides a basis for the construction of scaling tests using the means of mixtures of different sizes.

Variances and covariances among mixture total and component yields can be expressed in terms of the competition model. For the purpose of examining selection response, the following relations are necessary:

$$\begin{aligned}
\sigma_{v(m)v(k)} &= (\mathrm{var}\, c_{i/i} + (m_1 + k_1)\,\mathrm{cov}\, c_{i/i}\, c_{i/.} \\
&\quad + m_1\, k_1\, \mathrm{var}\, c_{i/.})/m^2 k^2 \\
\sigma_{a(m)a(k)} &= \mathrm{var}\, c_{./i}/m^2 k^2 \\
\sigma_{v(m)a(k)} &= (\mathrm{cov}\, c_{i/i}\, c_{./i} + m_1\,\mathrm{cov}\, c_{i/.}\, c_{./i})/m^2 k^2 \\
\sigma_{va(m)va(k)} &= \mathrm{var}\, c_{i/j}/m^2 k^2
\end{aligned}$$

where the variances and covariances of competition effects are defined with respect to the whole of the reference population, and var $c_{i/j}$ is an abbreviation for the variance of the residuals $(c_{i/j} - c_{i/.} - c_{./j} + c_{./.})$. Other statistics can be obtained by appropriate substitutions for k and m.

It is clear from these relations that when the competition model is adequate, then the a_j and v a_{ij} effects for mixtures of one size are perfectly correlated with the respective sets for mixtures of a different size. It can be shown that the correlation of v_i effects for mixtures of one size with those of another decreases monotonically as the difference in size is increased.

3 Selection response

3.1 Plant breeding context

A typical breeding programme involves a series of cycles of selection and reproduction during which the number of units is initially large but later is reduced as the final candidates for varietal production are identified. Because of the large numbers, a redefinition of the objective of the programme to produce mixture components rather than monocultures without seriously increasing the resources required means that each candidate can only be included once in a mixture with other candidates which are generally chosen at random. This is the principle assumed in the methods of mixture and component assessment to be investigated here. In general the number of components in the mixtures assessed (k) need not equal that of the mixtures to be finally produced (m). Only in the final choice of components for mixtures to be released will more extensive methods of testing be feasible, such as the production of all possible mixtures of given size among the candidates, or the inclusion of each in sufficient combinations to allow the estimation of general and specific effects as a basis for choosing certain single genotypes, pairs, or larger groups (group testing).

A second principle distinguishes these two phases in many programmes. If the candidates for selection are a group of established varieties or the final products of a breeding programme, then any mixtures which are constituted from them are ready for agricultural testing and subsequent release, and all gains, whether due to main effects of components or interactions among them, are utilised. The same will be true in a pedigree breeding programme in which the progenies at each generation can be assembled into the same combinations as those in which their parents were assessed for selection. If, on the other hand, selection is carried out on a population of families or genotypes at an intermediate stage of a recurrent selection programme, then the selected components will be assembled in a common pool for intercrossing so that many of the associations tested by the selection of components in mixtures will be dissolved, and the response due to the interaction of two or more components will be reduced. For most practical purposes it may be assumed that all the gains due to component interactions are lost following pooling. In fact, these gains will be further reduced by the intercrossing process itself, or by selfing in the pedigree programme, so that only the additive genetic component is retained, but questions about the inheritance of the various competition parameters must be considered as a separate, but extremely important topic. There are thus two separate situations with distinct consequences for the relative efficiencies of different selection methods.

3.2 General principles of the formulation of expected responses

The formulation of the expected response to selection among whole mixtures or their component genotypes is achieved by treating response in one variable (z) as a correlated effect of selection imposed on another (w). Then, according to the usual principles of linear prediction

$$R(z/w) = i \operatorname{cov}(z, w)/\operatorname{var}(w)^{1/2}$$

where cov(z, w) is the covariance of the true values of z and w, var(w) is the variance of the observed values of w, and i is the selection intensity in standard deviation units (Falconer 1960). This gain is the difference in mean value of mixtures of size m constructed from components drawn from the unselected and the selected populations. It should be noted that the selection intensity when is a function only of the numbers of candidates assessed and selected, and not of the size of mixtures used, it does not influence comparisons among different methods.

The expectation of any covariance required for the formulation of the response to selection of a particular type can be found directly from model (1). However, more general expressions can be derived which emphasise the overall form and properties of such statistics and allow specific cases to be extracted as required. A very general covariance can be defined as that between l components of a mixture of k components and q components of a mixture of m, the mixtures having a total of u components in common, t of which also occur in the group of q, s in the group of l, and r are common to both groups. This definition therefore depends on eight parameters, although only six of these determine the coefficients of the components of covariance, k and m serving to define these components. In terms of model (1), and including terms up to the first order of interaction

$$\operatorname{cov}^{(r,s,t,u;)}_{(l,q,k,m)} = r\,\sigma_{v(m)v(k)} + r\,l_1\,\sigma_{v(m)a(k)} + r\,q_1\,\sigma_{a(m)v(k)} + (l\,q\,u - q\,t - l\,s + r)\,\sigma_{a(m)a(k)} + r\,u_1\,\sigma_{va(m)va(k)}.$$

A simpler form of this covariance which is adequate for the description of most selection problems applies to the case in which each group contains only components which are either common to the other group or are unique (i.e. s = t = r), so that, with six parameters

$$\begin{aligned}\operatorname{cov}(r, u; l, q, k, m) = {} & r\,\sigma_{v(m)v(k)} + r\,l_1\,\sigma_{v(m)a(k)} \\ & + r\,q_1\,\sigma_{a(m)v(k)} + (r\,l_1\,q_1 \\ & + (u - r)lq)\,\sigma_{a(m)a(k)} \\ & + r\,u_1\,\sigma_{va(m)va(k)}.\end{aligned} \quad (4)$$

The expectations of cov(z, w) and var(w) for particular selection methods can be found by substitution of the appropriate values for the six parameters in (4) above, although var(w) will also contain components of error variance. However, when the selection unit is a group of p components (i.e. p will equal k for selection among whole mixtures, and 1 for individual component selection), then the covariance has to be multiplied by m/p to give cov(z, w), since this is the number of groups required to make up one whole mixture of m.

The values of the six parameters r, u, l, q, k, and m in cov(z, w) and var(w) for various types of selection are given in Table 1, together with their resultant expectations in terms of model (1). Those for var(w) show some of the contributions from higher order effects as well as from σ^2_{ew} and σ^2_{eb}, the error variances arising within and between mixture plots. Table 2 gives the same expectations (excluding the error components) in terms of the competition model.

Each method of selection can be characterised in terms of the selection criterion which it effectively applies to each candidate. These can be deduced from Tables 1 and 2, because the criterion is the term w in both cov(z, w) and var(w) which gives the expanded form in terms of the main effects of components. The response criterion is similarly z, and is always proportional to $v_{(m)} + m_1\,a_{(m)}$, or to $c_{i/i} + m_1(c_{i/.} + c_{./i})$ under the competition model.

3.3 Between mixture selection

When the components of a mixture cannot be separately harvested, then selection can only be carried out on the basis of the total yields of whole mixtures,

Table 1. The expectations of cov(z,w) and var(w) for different selection methods expressed in terms of a statistical model

	cov(z,w)	Common factor[a]	$\sigma_{v(m)v(k)}$	$\sigma_{v(m)a(k)}$	$\sigma_{a(m)v(k)}$	$\sigma_{a(m)a(k)}$	$\sigma_{va(m)va(k)}$
Between mixtures	m/k cov (k, k; k, m, k, m)	m	1	k_1	m_1	m_1k_1	k_1
Unrestricted:							
individual	m cov (1, 1; 1, m, k, m)	m	1	0	m_1	0	0
group	m/p cov (p, p; p, m, k, m)	m	1	p_1	m_1	m_1p_1	p_1
Within mixtures:							
individual		mk_1/k	1	−1	m_1	$-m_1$	0
group		m(k-p)/k	1	−1	m_1	$-m_1$	p_1
Group testing	m/p cov (p, p; k, m, k, m)	m	1	k_1	m_1	k_1m_1	p_1

	var(w)	common factor[a]	$\sigma^2_{v(k)}$	$\sigma^2_{a(k)}$	$\sigma_{v(k)a(k)}$	$\sigma^2_{va(k)}$	$\sigma^2_{aa(k)}$	$\sigma^2_{vaa(k)}$	σ^2_{eb}	σ^2_{ew}
Between mixtures	cov(k, k; k, k, k, k)	k	1	k_1^2	$2k_1$	k_1	$\frac{1}{2}k_1k_2^2$	$\frac{1}{2}k_1k_2$	1	1/k
Unrestricted:										
individual	cov(1, k; 1, 1, k, k)	1	1	k_1	0	k_1	$\frac{1}{2}k_1k_2$	$\frac{1}{2}k_1k_2$	$1/k^2$	$1/k^2$
group	cov(p, k; p, p, k, k)	p	1	k_2p+1	$2p_1$	k_1	–	–	p^2/k^2	p/k^2
Within mixtures:										
individual		k_1/k	1	1	−2	k_1	k_2	$\frac{1}{2}k_1k_2$	0	$1/k^2$
group		p(k-p)/k	1	1	−2	k_1	–	–	0	p/k^2
Group testing	cov(p, p; k, k, k, k)	p	1	k_1^2	$2k_1$	p_1	$\frac{1}{2}p_1k_2^2$	$\frac{1}{2}p_1p_2$	1/N	1/kN

[a] To simplify the form of the coefficients, a common factor has been extracted from all element other than σ^2_{eb} and σ^2_{ew} in each row
– Coefficient not evaluated

Table 2. The expectations of cov(z, w) and var(w) for different selection methods expressed in terms of a competition model

cov(z, w)	Common factor[a]	var $c_{i/i}$	cov $c_{i/i}c_{i/.}$	cov $c_{i/i}c_{./i}$	cov $c_{i/.}c_{./i}$	var $c_{i/.}$	var $c_{./i}$	var $c_{i/j}$
Between mixtures	$1/mk^2$	1	m_1+k_1	m_1+k_1	$2m_1k_1$	m_1k_1	m_1k_1	k_1
Unrestricted:								
individual	$1/mk^2$	1	m_1+k_1	m_1	m_1k_1	m_1k_1	0	0
group	$1/mk^2$	1	m_1+k_1	m_1+p_1	$(p_1+k_1)m_1$	m_1k_1	m_1p_1	p_1
Within mixtures:								
individual	k_1/mk^3	1	m_1+k_1	m_2	m_1k_2	m_1k_1	$-m_1$	0
group	$(k-p)/mk^3$	1	m_1+k_1	m_2	m_1k_2	m_1k_1	$-m_1$	p_1
Group testing	$1/mk^2$	1	m_1+k_1	m_1+k_1	$2m_1k_2$	m_1k_1	m_1k_1	p_1
var(w)								
Between mixtures	$1/k^3$	1	$2k_1$	$2k_1$	$2k_1^2$	k_1^2	k_1^2	k_1
Unrestricted:								
individual	$1/k^4$	1	$2k_1$	0	0	k_1^2	k_1	k_1
group	p/k^4	1	$2k_1$	$2p_1$	$2p_1k_1$	k_1^2	k_2p+1	k_1
Within mixtures:								
individual	k_1/k^5	1	$2k_1$	−2	$-2k_1$	k_1^2	1	k_1
group	$p(k-p)/k^5$	1	$2k_1$	−2	$-2k_1$	k_1^2	1	k_1
Group testing	p/k^4	1	$2k_1$	$2k_1$	$2k_1^2$	k_1^2	k_1^2	p_1

[a] To simplify the form of the coefficients, a common factor has been extracted from all elements in each row

accepting or rejecting all k components together. If a population of candidate genotypes is assembled at random into mixtures, the effect of such grouping is to reduce the rate of response to selection, and in comparison with unrestricted individual component selection, Table 1 shows an inflation of var(w) leading to a reduction in response by main effects by a factor of $k^{1/2}$. It should be noted that no variances arising among components within the mixtures (other than error) contribute to var(w) because these components are the fixed constitution of the mixture, and not a sample.

One property which is apparent from Table 1 is the occurrence of covariances between direct (v) and associate (a) effects in both cov(z, w) and var(w). This is because, in terms of main effects, selection is practised for $v_{(k)} + k_1 a_{(k)}$ and response is measured in terms of $v_{(m)} + m_1 a_{(m)}$. When the competition model is adequate, these criteria can be written as $c_{i/i} + k_1 (c_{i/.} + c_{./i})$ and $c_{i/i} + m_1 (c_{i/.} + c_{./i})$. Response to mixture selection is always positive when $k = m$, and this is the only simple method of selection which can ensure this outcome, since cov(z, w) can be written as a variance (see also Griffing 1967). Although the statistical model can give no information as to the likelihood of negative responses when $k \neq m$, these are seen to be unlikely under the competition model unless the $c_{i/i}$ terms are negatively correlated with $(c_{i/.} + c_{./i})$ and m and k are widely different.

The major benefit of mixture selection is its utilisation of component interactions up to the kth order. This benefit is lost when pooling follows selection, and in any case there is no advantage in increasing k beyond m. Again, no relationship can be established between interaction parameters defined for mixtures of different sizes, and the possibility of negative contributions to response from interactions cannot be ruled out. However, under the competition model only interactions of the first order are possible, and these are perfectly correlated among mixtures of different sizes.

Selection could be based on random groups of whole mixtures, but as pointed out by Griffing (1968), this is unlikely to be efficient, and in fact causes a reduction in response of the same order as that induced by grouping of components (i.e. a factor of $p^{1/2}$ for groups of size p).

3.4 *Unrestricted component selection*

Unrestricted component selection involves the separate assessment of components and their selection on the basis of observed yields without any reference to the yields of the mixtures in which they occur. Table 1 shows that cov(z, w) takes a very simple form because the selection criterion is now dependent simply on $v_{(k)}$ effects (or $c_{i/i} + k_1 c_{i/.}$). Negative response can occur whenever $m_1 \sigma_{a(m)v(k)}$ is negative and exceeds $\sigma_{v(m)v(k)}$. The competition model shows this to be more likely than with between mixture selection, especially with large mixtures and when $c_{i/.}$ and $c_{./i}$ are negatively related, as will normally be the case.

Although individual component selection avoids the overall reductions in gain caused by grouping, it can make no use of component interactions. A compromise measure which might be considered is selection among groups of p components, each group formed from components of the same mixture. Grouping incurs the usual reduction in gain of $p^{1/2}$, and Table 1 also shows that the contribution of $\sigma_{a(m)a(k)}$ to var(w) in this case is greater than when whole mixtures of the same size are used. This is due to a contribution made by the $(k - p)$ components of the mixture not contained in the group itself. In spite of these factors, it is not possible to predict with any certainty that within-mixture grouping will be less efficient than whole mixture selection, because the selection criteria themselves differ, that for groups of p being $c_{i/i} + k_1 c_{i/.} + p_1 c_{./i}$, and that for whole mixtures of p being $c_{i/i} + p_1 (c_{i/.} + c_{./i})$.

3.5 *Within mixture selection*

In this case the yields of components are expressed as deviations from their mixture means before selection is applied. The expectations of cov(z, w) and var(w) are not found directly from the general covariance expression, but by difference between unrestricted and mixture statistics. The chief property of within-mixture selection not shared by the other methods is its freedom from between plot errors (σ^2_{eb}), and this could be a significant advantage in many situations. However, it carries an even greater risk of negative response than does unrestricted component selection, because direct and associate effects are in opposition in the selection criterion of $v_{(k)} - a_{(k)}$ (or $c_{i/i} + k_1 c_{i/.} - c_{./i}$), and the covariance of v and a effects makes a negative contribution. Response to selection will be negative whenever $\sigma_{v(m)a(k)} + m_1 \sigma_{a(m)a(k)}$ exceeds $\sigma_{v(m)v(k)} + m_1 \sigma_{a(m)v(k)}$, which for $m = k = 2$ simply depends on $\sigma^2_a > \sigma^2_v$. With very large mixtures, the expectations of unrestricted and within mixture selection are very similar, but with small mixtures, within mixture selection suffers an overall disadvantage of $(k_1/k)^{1/2}$. It is worth noting that when k is large and $m = 1$, the situation described is that in which individuals undergoing mixed competition are being selected for monoculture performance, as is the case for segregants in a normal pedigree system. Response is dependent entirely on $\mathrm{cov}\, c_{i/i} c_{i/.}$, that is the covariance of monoculture performance with direct effects defined under purely inter-component competition.

As in the case of unrestricted selection, components could be grouped for selection within mixtures, but now also evaluated on a within mixture basis. It is interesting that such grouping has no effect on the selection criterion, and does not lead to the inflation of var (w) due to associate effects which occurs with unrestricted selection. However, the overall reduction in expected response due to the combined effects of grouping and within mixture restriction is of the order of $((k-p)/k\,p)^{1/2}$, so that this method is not expected to be efficient.

It may be noted that between mixture, unrestricted and within mixture selection methods are all specific types of a selection index employing information on mixture means and within mixture deviations. It is possible that all of these are markedly inferior to the optimum index suggested by Griffing (1969), for which estimates of cov (z, w)/var (w) for between and within mixtures would be used as weights. The problem in practice would be to obtain estimates of these statistics.

3.6 *Group Testing*

When interest is centred on relatively few varieties or genotypes as final candidates for mixture production, then the evaluation and selection of each variety may be based on the average value of a group of mixtures in which the variety is a common constituent, the others being chosen randomly or systematically from the remainder of the population. In analogy with progeny testing in genetic selection, estimates of the main or general effects of candidates are sought. Accordingly this method will be referred to here as group testing.

If the group used is large, and contains no two mixtures with more than the one common constituent, then response depends directly on the covariance of two mixtures with one common component. In practice however, such groups of mixtures will be restricted in size and may contain sub-groups with two and successively higher numbers of common constituents. In this case,

$$\mathrm{var}\,(w) = 1/N\left(\sum_i^k (n_i - n_{i+1})\,\mathrm{cov}\,(i, i; k, m; k, m)\right) + \sigma_{eb}^2/N + \sigma_{ew}^2/N\,k$$

in which n_i is the size of the sub-group with i common constituents. If no such sub-groups exist, then the above statistic is only influenced by the limited group size, so that $n_1 = N$, $n_2 = n_3 = \ldots = n_k = 1$, and

$$\mathrm{var}\,(w) = (N-1)/N\,\mathrm{cov}\,(1, 1; k, m; k, m) + 1/N\,\mathrm{cov}\,(k, k; k, m; k, m) + \sigma_{eb}^2/N + \sigma_{ew}^2/N\,k\,.$$

In each case, cov (z, w) contains only the main effect components of covariance which occur in var (w), so that

$$\mathrm{cov}\,(z, w) = 1/N\left(\sum_i^k (n_i - n_{i+1})\,i\,\sigma_{g(m)\,g(k)}\right).$$

More generally, selection can be carried out among sets of p constituents on the basis of groups of mixtures to which they are common, so that summation in var (w) is taken over sub-groups with from p to k common constituents, and cov (z, w) contains components up to the p-th order (Table 1).

Simple between mixture selection is a special case of group testing with a group size of one (i.e. $N = n_i = 1$ for all i), and serves as an appropriate basis for the assessment of the effects of increasing group size. The chief consequence of using groups is the reduction or removal of all interactions of higher order than the size of the common set (i.e. p). Since the method is only likely to be considered late in a programme, and at this stage it is possible to utilise all important interactions as well as main effects, the only situation in which it could have any possible advantage over mixture selection is when all effects and interactions up to the pth order are large, and those of higher order are insignificant. In view of its inherent costs and necessarily lower selection intensity, it can be concluded that group testing is unlikely to play a major role in component selection programmes.

3.7 *Comparisons among methods*

When statistical analysis indicates the absence of all but main effects, then it might be assumed that monoculture assessment is an adequate means of evaluating performance in mixture. However, it has already been pointed out that even under the first-order model of competition, the main effects v_i and a_i for monocultures or mixtures of one size are not necessarily strongly correlated with those of another, and this property is unaffected by the absence of first-order terms (i.e. $\mathrm{var}\,c_{i/j} = 0$). Even this simplest situation is governed by six parameters (Table 2) and so leads to very few general statements as to the superiority of any type of selection.

However, the conditions governing one particular comparison, that between the efficiencies of assessment in monoculture and that in mixtures of size m (i.e. $k = m$), can be formulated more exactly. Initially assuming error variance to be zero, the ratio of the expected gains is

$$R\,(\mathrm{mono})/R\,(\mathrm{mix}) = (\mathrm{var}\,c_{i/i} + m_1\,\mathrm{cov}\,c_{i/i}\,c_i)/ (\mathrm{var}\,c_{i/i}\,(\mathrm{var}\,c_{i/i} + 2\,m_1\,\mathrm{cov}\,c_{i/i}\,c_i + m_1^2\,\mathrm{var}\,c_i)/m)^{1/2}$$

where c_i is written for $(c_{i/.} + c_{./i})$. Monoculture assessment is the more efficient when this ratio exceeds unity, and this depends on

$$m_l (m r^2 - 1) + 2 m_l r t^{1/2} + t > 0$$

where r is the correlation of $c_{i/i}$ and c_i effects and t is the ratio var $c_{i/i}$/var c_i. The efficiency of monoculture assessment relative to that of mixtures is enhanced as either r or t is increased, so that it is the greater when r is positive and t exceeds m_l or with any value of r when t exceeds $(m\, m_l)$. In the great majority of competitive situations, the high yield of one component is achieved at the expense of, rather than to the benefit of its associates, so that the true inter-component direct and associate effects $c_{i/.}$ and $c_{./i}$ are strongly negatively correlated. This reduces the variance of their sum c_i which may therefore be expected to be lower than the monoculture variance, so that t is greater than unity. The sign of the correlation r depends on whether it is the more or the less aggressive competitors which tend to have the higher yields in monoculture, and although there are recorded instances in which high yield is associated with low aggressivity (Hamblin and Donald 1974), more commonly r will be positive. Because of their larger variance, the efficiency of monoculture assessment is less severely affected by error variation than that of mixtures, so that there are clearly many situations in practice, particularly for small values of m, when monoculture assessment is to be preferred to mixtures.

This conclusion may be checked, and some other comparisons made, by computing the expected efficiencies of different types of selection over a range of parameter values and typical mixture sizes. To do this, the six parameters in Table 2 were reduced to three by arbitrarily fixing the value of var $c_{i/i}$, and by assuming $c_{i/.}$ and $c_{./i}$ to be perfectly negatively correlated and $r\, c_{i/i}\, c_{./i}$ to be equal in magnitude and opposite in sign. Responses to selection among monocultures, to unrestricted selection, and selection between and within mixtures of two, three and four components were evaluated for a range of models with the ratio $t = \text{var}\, c_{i/i}/\text{var}\, c_{i/.}$ varying between 1/16 and 16, with var $c_{./i} \leqq$ var $c_{i/.}$, and $r\, c_{i/i}\, c_{i/.}$ from -1 to 1, again ignoring error variation. When the objective was to select components for binary mixtures, the highest responses were predicted for monoculture selection in 59% of cases and for unrestricted selection in the remainder, the response for unrestricted selection varying only slightly with changes in mixture size. Within mixture selection was always inferior due to the small size of mixture. If it is assumed that components could not be separately measured, so that unrestricted and within mixture selection could not be carried out, then monocultures were preferred in 69% of cases and binary mixtures for the remainder. Similar computations for the improvement of tertiary mixtures gave figures of 47 : 53 for monocultures and unrestricted selection when components could be separated, and 57 : 43 for monocultures and binary mixtures when they could not. If only the more likely set of parameter values with t greater than one are considered, the comparative figures for binary mixture improvement become 27 : 3 and 29 : 1, and for tertiaries 19 : 11 and 25 : 5. In all cases, comparisons of the methods on the basis of the incidence of negative responses gave a similar pattern.

Tables 1 and 2 show that the existence of any form of error variation will favour unrestricted and monoculture selection over mixtures, a differential effect on these two depending on the balance of within and between plot errors. It may therefore be concluded that under a competition model leading only to main effects, monoculture selection has at least as much potential as unrestricted selection for the improvement of binary or tertiary mixtures, both being superior to between mixture selection.

4 Conclusions

In practice, the total gain to be expected from any method of selection when the target and assessed mixtures are of the same size $(k = m)$ can be predicted from the true variance of selection units, whether components or mixtures, and the standard deviation of their observed values, so that separation of treatment and error variance components is all that is necessary. When $k \neq m$, then mixtures of the target size will also need to be grown so as to allow analysis of the covariation of the two types of mixture and estimation of the true covariance of the selection units. In neither case is a knowledge of the component structure of variances and covariances, as provided by Table 1, necessary for prediction. Only if a prediction of the loss of gain on pooling of components were required would a more complex analysis of structured sets of mixtures be needed to separate the covariance due to component interaction from that of main effects.

However, the main purpose of this paper is to identify the essential properties of the various selection methods, and as far as possible make general predictions as their relative efficiencies. The following important principles have been recognised.

(i) Selection among groups, whether these are whole mixtures or groups of components from them, allows evaluation and selection of component interactions, but at the expense of expressed variation and overall response.

(ii) The overall and relative magnitudes of error variation arising within and between plots influences the relative efficiencies of the different methods. For instance, within mixture selection may become an attractive prospect when the variation among plots is large.

(iii) Positive response to selection can only be assured when selection is practised among mixtures of the required size. Individual component selection can result in negative response whatever mixture size is used, particularly if assessed on a within mixture basis.

(iv) When components are pooled after selection, as will commonly be the case in the intermediate stages of breeding programmes, then little response due to interactions of components can be retained. Thus individual component selection will have a relatively higher efficiency than when mixtures are being finally constructed.

(v) When the competition model holds, the correlation between effects defined for mixtures of different sizes is a decreasing function of this difference. The overall effect of the size of mixture used for between mixture selection is therefore determined jointly by this correlation and the deleterious effects of grouping.

Hence monoculture selection has been shown to be very effective for binary mixture improvement, a little less so for tertiaries, and may be relatively inefficient for the improvement of large mixtures.

References

Falconer DA (1960) Introduction to quantitative genetics. Oliver and Boyd, London

Griffing B (1967) Selection in reference to biological groups. 1. Individual and group selection applied to populations of unordered groups. Aust J Biol Sci 20:127–139

Griffing B (1968) Selection in reference to biological groups. 2. Consequences of selection in groups of one size when evaluated in groups of a different size. Aust J Biol Sci 21: 163–170

Griffing B (1969) Selection in reference to biological groups. 4. Application of selection index theory. Aust J Biol Sci 22: 131–142

Hamblin J, Donald CM (1974) The relationships between plant form, competitive ability and grain yield in a barley cross. Euphytica 23:519–526

Wright AJ (1982) Some implications of a first-order model of interplant competition for the means and variances of complex mixtures. Theor Appl Genet 64:91–96