THE EFFECT OF FIELD BLOCKING
ON GAIN FROM SELECTION

by

W. M. Schutz and C. C. Cockerham

Institute of Statistics
Mimeograph Series No. 328
August, 1962

1.0 TABLE OF CONTENTS

	Page
2.0 LIST OF TABLES	v
3.0 LIST OF FIGURES	vii
4.0 INTRODUCTION	1
5.0 GENERAL FORMULATION OF THE EXPECTED GAIN FROM SELECTION. . . .	3
6.0 RELATIONSHIP OF THE WITHIN BLOCK VARIANCE TO BLOCK SIZE. . . .	6
7.0 THE BLOCK IN REPLICATION (BIR) DESIGN	10
7.1 Model and Analysis of Variance	10
7.2 General Formulation of the Prediction Equation	13
7.3 Average Error Variances of Entry Comparisons	16
7.4 Expected Gain from Selection	19
7.5 Comparisons of Methods of Selection	26
7.6 Effect of Replication of Entries on Expected Gain from Selection	33
7.7 Inclusion of Family Structure in Selection Schemes Involving the BIR Design	39
8.0 THE REPLICATION IN BLOCKS (RIB) DESIGN	48
8.1 Model and Analysis of Variance	48
8.2 Average Error Variance of Entry Comparisons	51
8.3 Expected Gain from Selection	54
8.4 Comparisons of Methods of Selection	56
8.5 Effect of Replication of Entries on Expected Gain from Selection	60
9.0 THE SIMPLE LATTICE DESIGN	62
9.1 Model and Equations for Adjusted Entry Means	62
9.2 Average Error Variance of Entry Comparisons	67
9.3 Expected Gain from Selection	69
9.4 Comparisons of Methods of Selection	74
10.0 GENERAL DISCUSSION AND CONCLUSIONS	78
11.0 SUMMARY	85
12.0 LIST OF REFERENCES	86

2.0 LIST OF TABLES

Page

1. Analysis of the environmental variance in a field of g groups with n plots per group	8
2. Analysis of variance for the BIR design with expectations of mean squares for Models 7.1.1 and 7.1.2	10
3. Analysis of variance and expectations of mean squares for the BIR design based on Model 7.1.3	11
4. Total number of pair-wise entry comparisons between and within blocks when checks are included in the BIR design . .	17
5. Ratio of the expected gain from within blocks selection to the expected gain from selection ignoring blocks, Δ_1/Δ_2 , for $T=256$, $r=2$, and σ_p^2/σ_T^2 , approaching zero	30
6. Ratio of the expected gain from within blocks selection to the expected gain from selection ignoring blocks, Δ_1/Δ_2 , for $T=256$, $r=2$, and $\sigma_p^2/\sigma_T^2 = 0.5$	30
7. Ratio of the expected gain from selection using the best index to the expected gain from selection ignoring blocks, Δ_1/Δ_2 , for $T=256$, $r=2$, and σ_p^2/σ_T^2 , approaching zero	34
8. Ratio of the expected gain from selection using the best index to the expected gain from selection ignoring blocks, Δ_1/Δ_2 , for $T=256$, $r=2$, and $\sigma_p^2/\sigma_T^2 = 0.5$	34
9. Ratio of the expected gain from selection among entry means adjusted for check deviations to the expected gain from selection without checks, Δ_{2c}/Δ_2 , for $T=256$, $r=2$, and σ_p^2/σ_T^2 , approaching zero; when the top 25 entries are selected	35
10. Ratio of the expected gain from selection among entry means adjusted for check deviations to the expected gain from selection without checks, Δ_{2c}/Δ_2 , for $T=256$, $r=2$, and $\sigma_p^2/\sigma_T^2 = 0.5$; when the top 25 entries are selected	36
11. Analysis of variance for the RIB design with expectations of mean squares for Model 8.1.1 and Model 8.1.2	48

LIST OF TABLES (continued)

Page

12. Total number of pair-wise entry comparisons between and within blocks of the RIB design and the variance of each of these comparisons 51
13. Ratio of the average error variance per comparison of unadjusted treatment means to the total variance, $\bar{\sigma}_2^2/\sigma_T^2$, for $T=256$ and $r=2$ 53
14. Ratio of the expected gain from selection ignoring blocks for the BIR design to the expected gain for the RIB design, Δ_2/Δ_1 , for $T=256$, $r=2$, and σ_p^2/σ_T^2 approaching zero . 58
15. Ratio of the expected gain from selection ignoring blocks for the BIR design to the expected gain for the RIB design, Δ_2/Δ_1 , for $T=256$, $r=2$, and $\sigma_p^2/\sigma_T^2 = 0.5$ 58
16. Ratio of the expected gain from selection on the best index for the BIR design to the expected gain for the RIB design, Δ_1/Δ_T , for $T=256$, $r=2$, and σ_p^2/σ_T^2 approaching zero 59
17. Ratio of the expected gain from selection on the best index for the BIR design to the expected gain for the RIB design, Δ_1/Δ_T , for $T=256$, $r=2$, and $\sigma_p^2/\sigma_T^2 = 0.5$ 59
18. Values for the ratio Δ_2/Δ_1 , for an experiment of 512 plots, where Δ_2 is the expected gain from selection ignoring blocks with 2 replications and Δ_1 is the expected gain using 4 replications 61

3.0 LIST OF FIGURES

Page

1. Graph of the equation $\Delta_1 - \Delta_2 = 0$ for $T=256$, $r=2$, and σ_p^2/σ_T^2 , approaching zero 32
2. Graph of the equation $\Delta_1 - \Delta_2 = 0$ for $T=256$, $r=2$, and $\sigma_p^2/\sigma_T^2 = 0.5$ 32
3. Graph of the ratio Δ_1/Δ_1 for an experiment of 512 plots with $b=16$ and $B=0.5$, where Δ_1 is the expected gain from selection within blocks using 2 replications and Δ_1 is the expected gain using 4 replications 38
4. Graphical representation of the expected gain from selection for $T=64$, $r=2$, $\sigma_s^2/\sigma_T^2 = \sigma_d^2/\sigma_T^2 = 0.1$, and $B=0.5$, where Δ_{1f} = expected gain from selecting within blocks on full sib family means, Δ_{2f} = expected gain from selection ignoring blocks, and Δ_{If} = expected gain from selection using an index 46

4.0 INTRODUCTION

The selection of top performing genetic material on the basis of field tests has been practiced by plant breeders for many years. The choice of an environmental design in which to grow this material has been largely a matter of personal preference on the part of the breeder. In general, statisticians have evaluated the efficiency of experimental designs on the basis of the average error variances of treatments or entries, but it has not been clear whether this criterion is similar to that of gains expected from selecting among entries.

Perhaps the simplest design used by plant breeders is the randomized complete block design in which all entries appear together in the same block. If the experiment consists of a large number of entries, then a question arises as to the advisability of dividing the material into sub-groups or incomplete blocks in hopes of reducing the environmental variability.

If blocking is done, then one means of selection is to adjust the entry means for differences between blocks. However, in terms of true differences, superior entries which by chance fall into the same block will receive no credit for this but will only be compared among themselves. On the other hand, if blocking is ignored, which is the original alternative of randomized complete blocks, the entries in the higher incomplete blocks will often be high because of environmental advantages.

If one or more check plots or controls are included in each block, then a third possibility is selection on entry means adjusted for check performance.

A fourth alternative is an index which takes all of these sources of information into account and assigns appropriate weights to them.

The purpose of this study is to evaluate the effects of blocking on each of these methods of selection in terms of expected genetic gain and average error variance.

5.0 GENERAL FORMULATION OF THE EXPECTED GAIN FROM SELECTION

If the entries referred to in Section 4.0 are ranked according to their observed or predicted values, $\hat{Y}$'s, and the top h fraction is chosen, the expected gain from selection, Δ , is the expected difference between the mean of the true values, p 's, of selected entries and the mean of all entries. It should be emphasized that these entries may consist of lines, varieties, related or unrelated progenies, or any other kind of genetic material and the $\hat{Y}$'s involved may be the observed values of the entries themselves or predicted values based on a linear combination of factors as illustrated in Section 7.2. Smith (1936) showed that if the p 's and $\hat{Y}$'s are normally distributed, Δ may be expressed as a function of two quantities:

$$\Delta = s_d B_{p\hat{Y}}.$$

The quantity $B_{p\hat{Y}}$ is the regression of the true values of the entries on the $\hat{Y}$'s. The other quantity, s_d , is the selection differential, which is the expected mean difference in the $\hat{Y}$'s of the selected entries and all entries, that is

$$s_d = E(\hat{Y}_s - \hat{Y}).$$

The $\hat{Y}$'s may be transformed to a variate, v , with unit variance and mean at zero:

$$v = \frac{\hat{Y} - u_Y}{\sigma_{\hat{Y}}}$$

where u_Y and σ_Y are the population mean and standard deviation, respectively.

The mean value of p associated with any given value of $\hat{Y}$ is

$$E(p-u_p) = B_{p\hat{Y}}(\hat{Y}-u_Y) = v B_{p\hat{Y}} \sigma_{\hat{Y}}.$$

If the point of truncation separating the selected group from the discarded group is designated as v' , then summing for all values of v greater than v' and dividing by their frequency, h , results in

$$\Delta = E(\bar{p}_s - u_p) = \frac{1}{h} B_{p\hat{Y}} \sigma_{\hat{Y}} \int_{v'}^{\infty} v \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2} dv = \frac{z}{h} \sigma_{\hat{Y}} B_{p\hat{Y}} \quad (5.0.1)$$

where z is the ordinate of the unit normal curve at the deviate v' .

In practice, however, the upper fraction, h , of the sample is saved so that

$$\Delta = E(\bar{p}_s - u_p) = B_{p\hat{Y}} \sigma_{\hat{Y}} E(\bar{v}_s).$$

$E(\bar{v}_s)$ is an order statistic and depends on the size of the sample. For normally distributed samples under 50,

$$E(\bar{v}_s) = S_n$$

where S_n is computed from Table XX of Fisher and Yates (1953) for the appropriate sample size n , and

$$s_d = S_n \sigma_{\hat{Y}}.$$

For samples larger than 50 the large sample theory is of sufficient accuracy and

$$s_d \doteq S_{\infty} \sigma_{\hat{Y}}$$

where $S_{\infty} = \frac{z}{h}$ from equation 5.0.1.

Another approximation technique, and one which is used in this study, is

$$s_d \doteq S_{\infty} \sigma_{\hat{Y}_n}$$

where $\sigma_{\hat{Y}_n}^2 = E(\hat{Y} - \bar{\hat{Y}})^2$ and $\bar{\hat{Y}}$ is the sample mean of the $\hat{Y}$'s.

Expressing B_{pY} in an alternative form,

$$B_{pY} = \frac{\text{Cov } p\hat{Y}}{\sigma_{\hat{Y}}^2} = \frac{\text{Cov } p\hat{Y}_n}{\sigma_{\hat{Y}_n}^2}$$

results in a more common form of the expected gain from selection:

$$\Delta = \frac{S_{\infty} \text{Cov } p\hat{Y}_n}{\sigma_{\hat{Y}_n}^2}.$$

6.0 RELATIONSHIP OF THE WITHIN BLOCK VARIANCE TO BLOCK SIZE

Since the primary purpose of this study is to compare the effects of blocking on the progress from selection, it is necessary to relate the environmental components of variance to alternative blocking patterns. One means of doing this is by making use of an empirical relationship derived by Smith (1938).

Using data from a uniformity experiment with wheat, Smith found that the regression of the logarithms of the variance for plots of different areas on the logarithms of their areas was approximately linear (i.e., the relative reduction in variability of plot means for a relative increase in plot size is similar throughout the range observed). A graphical review of variances reported in the literature for 39 other uniformity experiments indicated that the results of most experiments conform to the same law.

In equational form, the variance among plots within a block may be expressed as

$$\log (V_x)_{n'} = \log (V_1)_n - B' \log x \quad (6.0.1)$$

where

$n' = n/x =$ number of plots per block

$n =$ size of the block

$(V_x)_{n'} =$ variance per unit area among n' plots of size x

$(V_1)_n =$ variance among plots of size unity in a block of size n

$B' =$ regression coefficient .

It should be pointed out that $(V_1)_n$ and B' can be estimated from data. The limiting values of the regression coefficient are zero and one unless inter-experimental unit competition is present. If plots are chosen at random to make up each block, the regression coefficient would be unity. However, where contiguous plots are included in a block, there is likely to be a positive correlation among these plots and B' tends to be less than one. A B' of zero implies a perfect correlation between contiguous plots.

Smith computed B' values for thirty-nine different sets of uniformity trial data and found that most of the regression coefficients fell within the range 0.2 to 0.8.

Smith pointed out that in the variance law, 6.0.1, the size of the field, n , is arbitrarily fixed and that if n is varied the regression becomes curved for any value other than that originally assigned. The regression is therefore inconsistent with the requirement that the law shall be unaffected by variation in the size of the field.

This difficulty is overcome if an infinitely large field is postulated in which an observed field represents a single block. Under these conditions, the variance law becomes:

$$\log (V_x)_{\infty} = \log (V_1)_{\infty} - B \log x . \quad (6.0.2)$$

The modified regression coefficient, B , is applicable to an infinite field and in general is not equal to B' .

Since blocks are in reality large plots, the variance between blocks may be estimated in the same way, i.e.,

$$\log (V_{xn'})_{\infty} = \log (V_1)_{\infty} - B \log xn' . \quad (6.0.3)$$

No generality is lost by setting x equal to one in subsequent developments when it represents the smallest unit or plot usual for the experiment considered. Under this condition, $n' = n$ and 6.0.3 becomes:

$$\log (V_n)_{\infty} = \log (V_1)_{\infty} - B \log n . \quad (6.0.4)$$

If a field is subdivided into g groups or blocks with n plots per group, the analysis of environmental variance may be represented as in Table 1.

Table 1. Analysis of the environmental variance in a field of g groups with n plots per group

Source of variation	d.f.	Mean squares
Between groups	$g-1$	$n(V_n)_g$
Within groups	$g(n-1)$	$(V_1)_n$
Total	$gn-1$	$(V_1)_{gn}$

From Table 1 it follows that the variance within groups of n plots is

$$(V_1)_n = \frac{(gn-1)(V_1)_{gn} - n(g-1)(V_n)_g}{g(n-1)} . \quad (6.0.5)$$

To relate $(V_1)_n$ to an infinite field, it is necessary for g to take the value of infinity in 6.0.5 yielding:

$$\begin{aligned} (V_1)_n &= \frac{n(V_1)_{\infty} - n(V_n)_{\infty}}{n-1} \\ &= \frac{n(1-n^{-B})(V_1)_{\infty}}{n-1} . \end{aligned} \quad (6.0.6)$$

For the purposes of this study, it is of interest to employ 6.0.6 to express the variance within blocks of p plots,

$$(V_1)_p = \frac{p(1-p^{-B})(V_1)_\infty}{p-1}, \quad (6.0.7)$$

the variance within experiments of $T = bp$ plots,

$$(V_1)_T = \frac{T(1-T^{-B})(V_1)_\infty}{T-1}, \quad (6.0.8)$$

and the variance within replicated experiments of $T' = rbp$ plots,

$$(V_1)_{T'} = \frac{T'[1-T']^{-B}(V_1)_\infty}{T'-1}. \quad (6.0.9)$$

Use is made of these variance relationships in later sections in writing the environmental components of variance for specific designs in terms of p , T , T' , B , and $(V_1)_\infty$.

7.0 THE BLOCK IN REPLICATION (BIR) DESIGN

7.1 Model and Analysis of Variance

Consider an experiment in which each block within a replication contains a random group of entries and in which the blocks are allocated at random to an area in the replication. A model for such an experiment is as follows:

$$Y_{ijk} = u + B_i + r_k + p_{ij} + b_{ki} + (pr)_{ijk} \quad (7.1.1)$$

The effects are B for blocks, r for replications, p for entries in blocks, b for environmental differences of blocks in replications, and pr for replication x entry in block interactions. The subscript i varies from one to b, j varies from one to p, and k varies from one to r. All effects are assumed to be random and uncorrelated which leads to the expectations of mean squares in the analysis of variance given in Table 2.

Table 2. Analysis of variance for the BIR design with expectations of mean squares for Models 7.1.1 and 7.1.2

Source of variation	d.f.	Expectations of mean squares	
		Model 7.1.1	Model 7.1.2
Replications	r-1	$\sigma^2 + p\sigma_b^2 + b p \sigma_r^2$	$\sigma^2 + p\sigma_b^2 + b p \sigma_r^2$
Blocks	b-1	$\sigma^2 + r\sigma_p^2 + p\sigma_b^2 + pr\sigma_B^2$	$\sigma^2 + p\sigma_b^2$
Blocks x reps	(r-1)(b-1)	$\sigma^2 + p\sigma_b^2$	$\sigma^2 + p\sigma_b^2$
Entries in blocks	b(p-1)	$\sigma^2 + r\sigma_p^2$	σ^2
Reps x entries in bl.	b(p-1)(r-1)	σ^2	σ^2

To further clarify the components of environmental variance involved, consider the entries to be uniform. The model then becomes:

$$Y_{ijk} = u + r_k + b_{ki} + e_{ijk} . \quad (7.1.2)$$

The effects are r for replications, b for blocks in replications, and e for plots. The subscripts are the same as for Model 7.1.1. Model 7.1.2 leads to the expectations of mean squares given in Table 2.

Note that under the assumption of uniformity, there is no longer any reason for grouping blocks together from different replications, except at random. Thus, the effect designated as B_i in Model 7.1.1 is not, in reality, a valid reflection of environmental differences among blocks and the proper model for the experiment is

$$Y_{ijk} = u + r_k + b_{ki} + p_{ij} + (pr)_{ijk} . \quad (7.1.3)$$

The effects are r for replications, b for blocks in replications, p for entries in blocks, and pr for replication x entry in block interactions, which leads to the analysis of variance presented in Table 3.

Table 3. Analysis of variance and expectations of mean squares for the BIR design based on Model 7.1.3

Source of variation	d.f.	Expectations of mean squares
Replications	$r-1$	$\sigma^2 + p\sigma_b^2 + bp\sigma_r^2$
Blocks in reps	$r(b-1)$	$\sigma^2 + p\sigma_b^2 + \sigma_p^2$
Entries in blocks	$b(p-1)$	$\sigma^2 + r\sigma_p^2$
Reps x entries in bl.	$b(p-1)(r-1)$	σ^2

If uniformity of entries is assumed, the component of variance for blocks may be derived from the analysis of variance in Table 3 in terms of the total variance within a replication, σ_T^2 , and the variance between plots within a block, σ^2 . That is,

$$\sigma_b^2 = \frac{b(T-1)}{T(b-1)} (\sigma_T^2 - \sigma^2) . \quad (7.1.4)$$

The quantity $\frac{b(T-1)}{T(b-1)}$ is a function of the number of blocks which is easily evaluated. If there is a positive correlation between plots within a block, σ^2 would be expected to decrease as block size decreases. A useful means of quantitizing this relationship is available from Smith's variance law as outlined in Section 6.0. An application of equations 6.0.7 and 6.0.8 permits the expression of σ^2 , σ_T^2 , and σ_b^2 in terms of T, B, b, and $(V_1)_\infty$.

For ease in computation, the components of variance are expressed as fractions of the total variance, $\sigma_{T'}^2$, among $T' = rbp$ plots. This eliminates the constant quantity $(V_1)_\infty$ and in no way affects the comparisons to be made in subsequent sections among various methods of selection, since $\sigma_{T'}^2$ is constant for an experiment of a given size.

Thus,

$$\frac{\sigma^2}{\sigma_{T'}^2} = \frac{T'-1}{r(T-b)} \left[\frac{(T')^B - (rb)^B}{(T')^B - 1} \right] \quad (7.1.5)$$

$$\frac{\sigma_b^2}{\sigma_{T'}^2} = \frac{b(T-1)}{T(b-1)} \left[\frac{\sigma_T^2}{\sigma_{T'}^2} - \frac{\sigma^2}{\sigma_{T'}^2} \right] \quad (7.1.6)$$

where

$$\frac{\sigma_T^2}{\sigma_{T'}^2} = \frac{T'-1}{r(T-1)} \left[\frac{(T')^B - r^B}{(T')^B - 1} \right] \quad (7.1.7)$$

and

$$T' = rbp = rT .$$

The limit of $\sigma^2/\sigma_{T'}^2$, as B approaches one is obviously one, but the limit as B approaches zero is not readily apparent and must be determined by methods of calculus.

$$\lim_{B \rightarrow 0} \frac{\sigma^2}{\sigma_{T'}^2} = \frac{T'-1}{r(T-b)} \left[1 - \frac{\log_e(br)}{\log_e T'} \right] .$$

Similarly, the limit of $\sigma_T^2/\sigma_{T'}^2$, is 1 as B approaches one and for B approaching zero is

$$\lim_{B \rightarrow 0} \frac{\sigma_T^2}{\sigma_{T'}^2} = \frac{T'-1}{r(T-1)} \left[1 - \frac{\log_e r}{\log_e T'} \right] .$$

If t check plots are included in each block as an additional source of information on environmental variability, the model for checks is

$$X_{ilk} = u_1 + b_{ki} + t_l + e_{ilk} . \quad (7.1.8)$$

The effects are r for replications, b for blocks in replications, t for checks, and e for plots. The subscript i varies from one to b, k varies from one to r, and l varies from one to t. All effects are assumed to be random and uncorrelated and in addition it is assumed that checks and entries react similarly to block and replication differences.

7.2 General Formulation of the Prediction Equation

A general formulation for the predicted value, $\hat{Y}_{ij}$, of the jth entry mean in the ith incomplete group which encompasses all of the

alternative methods of selection mentioned in Section 4.0 may be written as follows:

$$\hat{y}_{ij.} = y_{...} + \hat{\beta}_1(y_{ij.} - y_{i..}) + \hat{\beta}_2(y_{i..} - y_{...}) + \hat{\beta}_3(x_{i..} - x_{...}) \quad (7.2.1)$$

where Y_{ijk} is the value of the j th entry in the i th incomplete block in the k th replication and X_{ilk} is the value of the l th check in the i th incomplete block in the k th replication. The $\hat{\beta}$'s are weights to be assigned to the corresponding deviations in parentheses. Lower case x 's and y 's are used to indicate means and a dot indicates an average over that subscript. The first deviation is that of the entry mean from the block mean, the second deviation that of the block mean from the over-all mean, and the third deviation that of the block check mean from the over-all check mean.

For the method of selecting within incomplete blocks to be denoted as within blocks selection, the $\hat{\beta}$'s have the following values:

$$\hat{\beta}_1 = 1, \hat{\beta}_2 = \hat{\beta}_3 = 0$$

and the predicted value of an entry mean is

$$\hat{y}_{ij.} = y_{...} + (y_{ij.} - y_{i..}) \quad (7.2.2)$$

The second alternative method, which is to not utilize incomplete blocks, has the following values for the $\hat{\beta}$'s:

$$\hat{\beta}_1 = \hat{\beta}_2 = 1, \hat{\beta}_3 = 0$$

and

$$\hat{y}_{ij.} = y_{ij.} \quad (7.2.3)$$

If selection is based on entry means adjusted for check deviations, the $\hat{\beta}$'s are

$$\hat{\beta}_1 = \hat{\beta}_2 = 1, \hat{\beta}_3 = -1 .$$

Under these conditions, the predicted value of an entry mean is

$$\hat{y}_{ij.} = y_{ij.} - (x_{i..} - x_{...}) . \quad (7.2.4)$$

A method of selection somewhat comparable to within blocks selection may be obtained by adjusting entry means for the block deviations of both entries and checks. If these deviations are weighted in proportion to the number of entries and checks making up the block means, the $\hat{\beta}$'s are

$$\hat{\beta}_1 = 1, \hat{\beta}_2 = \frac{t}{p'+t}, \hat{\beta}_3 = \frac{-t}{p'+t}$$

where p' is the number of entries per block, t is the number of checks per block ($p=p'+t$), and

$$\hat{y}_{ij.} = y_{...} + (y_{ij.} - y_{i..}) + \frac{t}{p}(y_{i..} - y_{...}) - \frac{t}{p}(x_{i..} - x_{...}) . \quad (7.2.5)$$

For the best index method of selection, $\hat{\beta}$'s are chosen that will maximize the correlation between the predicted values of the entry means, $\hat{y}_{ij.}$'s, and the true values of the entry means. This procedure of giving some weight to the block deviation is an attempt to recover inter-block information. When check information is not included in the index, $\hat{\beta}_3$ is set equal to zero.

7.3 Average Error Variances of Entry Comparisons

The earliest work on the use of incomplete block designs in evaluating plant varieties was reported by Yates (1936) in which he outlined a series of designs called pseudo-factorials. This group of designs is now generally referred to as lattice designs and is used extensively in variety testing. Yates used the relative magnitudes of the average error variances of treatment comparisons as the criterion of efficiency and no attempt was or has been made to relate these variances to expected genetic gain.

If entry means are adjusted by utilizing the prediction equation, 7.2.1, the variance of within block comparisons is

$$\frac{2\hat{\beta}_1^2 \sigma^2}{r}$$

and the variance of between block comparisons is

$$\frac{2\hat{\beta}_1^2(p-1)\sigma^2}{p'r} + 2\hat{\beta}_2^2\left(\frac{\sigma_b^2}{r} + \frac{\sigma^2}{p'r}\right) + 2\hat{\beta}_3^2\left(\frac{\sigma_b^2}{r} + \frac{\sigma^2}{rt}\right) + 4\hat{\beta}_2\hat{\beta}_3\left(\frac{\sigma_b^2}{r}\right) + \frac{2(\hat{\beta}_2 - \hat{\beta}_1)^2(p-1)\sigma_p^2}{p^2}.$$

Note that when entry means are adjusted for block effects ($\hat{\beta}_2 \neq \hat{\beta}_1$), a pair-wise comparison of entries in different blocks involves entry effects other than the two being compared. These additional entry effects contribute to the error variance of the comparison.

The total number of pair-wise entry comparisons within and between blocks is presented in Table 4.

Table 4. Total number of pair-wise entry comparisons between and within blocks when checks are included in the BIR design.

Type of comparison	Total no.
Within blocks	$\frac{bp'(p'-1)}{2}$
Between blocks	$\frac{b(b-1)(p')^2}{2}$

When check plots are included in the experiment and entry means are adjusted for check deviations ($\hat{\beta}_1 = \hat{\beta}_2 = 1, \hat{\beta}_3 = -1$), the adjusted value of an entry mean is

$$\hat{y}_{ij.} = y_{...} + (y_{ij.} - y_{...}) - (x_{1..} - x_{...}) .$$

Under these conditions, the variance of within block comparisons is $\frac{2\sigma^2}{r}$ and the variance of between block comparisons is $\frac{2\sigma^2(t+1)}{rt}$. The average error variance per comparison is

$$\begin{aligned} \frac{\sigma^2}{2c} &= \frac{\frac{bp'(p'-1)\sigma^2}{r} + \frac{b(b-1)(p')^2(t+1)\sigma^2}{rt}}{\frac{bp'(bp'-1)}{2}} \\ &= \frac{2\sigma^2}{r} \left[\frac{p'(bt+b-1) - t}{t(bp'-1)} \right] . \end{aligned}$$

If entry means are adjusted for block deviations of both entries and checks as in equation 7.2.5, the variance of between block comparisons is

$$\frac{2(p-1)}{p} \left[\frac{\sigma^2}{r} + \frac{(p')^2 \sigma_p^2}{p^3} \right]$$

and the average error variance per comparison is

$$\bar{\sigma}_{1c}^2 = \left[\frac{b(p^2 - p - pt + t - t)}{p(bp - bt - 1)} \right] \frac{2\sigma^2}{r} + \frac{(b-1)(p-t)^3(p-1)2\sigma_p^2}{p^4(bp - bt - 1)} .$$

The relative magnitudes of $\bar{\sigma}_{2c}^2$ and $\bar{\sigma}_{1c}^2$ cannot be readily assessed since $\bar{\sigma}_{1c}^2$ is a function of both σ^2 and σ_p^2 . However, when σ_p^2 is small the variance of between block comparisons is smaller for the latter method of adjusting entry means and the within block variance is the same for both methods so that $\bar{\sigma}_{2c}^2 > \bar{\sigma}_{1c}^2$.

For the special case where $t = 0$ and $p = p'$, the adjusted value of an entry mean becomes:

$$\hat{y}_{1j.} = y_{...} + (y_{1j.} - y_{1..})$$

and the average error variance per comparison is

$$\bar{\sigma}_1^2 = \frac{T-b}{T-1} \left[\frac{2\sigma^2}{r} + \frac{(b-1)}{T} 2\sigma_p^2 \right] .$$

Note that as block size decreases, the coefficient of σ_p^2 increases so that the treatment contribution to the average error variance becomes relatively more important for small block sizes.

If $t = 0$ and block information is ignored (unadjusted entry means are compared) the average error variance per comparison is of course independent of any pseudo-blocking and is a function of σ_T^2 . The variance of between block comparisons becomes $\frac{2(\sigma^2 + \sigma_b^2)}{r}$, and the variance of within block comparisons is $\frac{2\sigma^2}{r}$ as before so that the average error variance per comparison is

$$\bar{\sigma}_2^2 = \frac{\frac{bp(p-1)}{r} \sigma^2 + \frac{b(b-1)p^2}{r} (\sigma^2 + \sigma_b^2)}{\frac{bp(bp-1)}{2}} = \frac{2}{r} \left[\sigma^2 + \frac{T(b-1)}{b(T-1)} \sigma_b^2 \right] .$$

If σ_b^2 is written as a function of σ^2 and σ_T^2 as in 7.1.4, then

$$\sigma_2^2 = \frac{2}{r} \left\{ \sigma^2 + \frac{T(b-1)}{b(T-1)} \left[\frac{b(T-1)}{T(b-1)} (\sigma_T^2 - \sigma^2) \right] \right\} = \left(\frac{2}{r} \right) \sigma_T^2.$$

Since σ_T^2 is constant for an experiment of a given size, σ_2^2 is independent of block size. Cochran and Cox (1957) showed that a similar statement can be made about the average error variance of comparisons between and within the whole plots of a split plot design. Therefore, the two designs are similar in that respect.

7.4 Expected Gain from Selection

In order to formulate the expected gain from selection for each of the proposed selection schemes, it is necessary to derive the variances and covariances of the deviations in the prediction equation, 7.2.1, and the covariances between the deviations and the true values of entry means. All derivations are based on Models 7.1.3 and 7.1.8.

The following expectations are utilized in obtaining the variances and covariances of the deviations:

$$E(x_{i..})(x_{...}) = E(x_{...})^2 = u^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{br} + \frac{\sigma_t^2}{t} + \frac{\sigma_{btr}^2}{btr}$$

$$E(y_{ij.})(y_{i..}) = E(y_{i..})^2 = u^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{r} + \frac{\sigma_p^2}{p} + \frac{\sigma_{rp}^2}{rp}$$

$$E(y_{ij.})(y_{...}) = E(y_{i..})(y_{...}) = E(y_{...})^2$$

$$= u^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{br} + \frac{\sigma_p^2}{bp} + \frac{\sigma_{bpr}^2}{bpr}$$

$$E(y_{...})(x_{i...}) = E(y_{ij.})(x_{...}) = E(y_{i..})(x_{...})$$

$$= E(y_{...})(x_{...}) = uu_1 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{br}$$

$$E(y_{ij.})(x_{i..}) = E(y_{i..})(x_{i..}) = uu_1 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{r}$$

$$E(y_{ij.})^2 = u^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{r} + \sigma_p^2 + \frac{\sigma^2}{r}$$

$$E(x_{i..})^2 = u_1^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{r} + \frac{\sigma_t^2}{t} + \frac{\sigma^2}{rt}$$

where $p' = \frac{T-bt}{b}$ and $p = p' + t$.

The variances and covariances of the deviations utilized in the alternative methods of selection are

$$V(y_{ij.} - y_{i..}) = E(y_{ij.})^2 - E(y_{i..})^2 = \frac{p'-1}{p'r}(\sigma^2 + r\sigma_p^2) \quad (7.4.1)$$

$$V(y_{i..} - y_{...}) = E(y_{i..})^2 - E(y_{...})^2 = \frac{b-1}{b}(\frac{\sigma_b^2}{r} + \frac{\sigma_p^2}{p'} + \frac{\sigma^2}{p'r}) \quad (7.4.2)$$

$$V(y_{ij.} - y_{...}) = E(y_{ij.})^2 - E(y_{...})^2 = \frac{(b-1)}{br}\sigma_b^2 + \frac{bp'-1}{bp'r}(\sigma^2 + r\sigma_p^2) \quad (7.4.3)$$

$$V(x_{i..} - x_{...}) = E(x_{i..})^2 - E(x_{...})^2 = \frac{b-1}{br}(\sigma_b^2 + \frac{\sigma^2}{t}) \quad (7.4.4)$$

$$\begin{aligned} V[(y_{ij.} - y_{...}) - (x_{i..} - x_{...})] &= E[(y_{ij.} - x_{i..}) - (y_{...} - x_{...})]^2 \\ &= \frac{(bp'-1)}{bp'}\sigma_p^2 + \frac{[p(bt+b-1)-t]}{bp'tr}\sigma^2 \end{aligned} \quad (7.4.5)$$

$$\begin{aligned}
& V[(y_{ij.} - y_{i..}) + \frac{t}{p}(y_{i..} - y_{...}) - \frac{t}{p}(x_{i..} - x_{...})] \\
& = E[(y_{ij.} - y_{i..}) + \frac{t}{p}(y_{i..} - y_{...}) - \frac{t}{p}(x_{i..} - x_{...})]^2 \\
& = \frac{[bp'(p^2 - 2p + p') - t^2]}{p^2 bp'} \sigma_p^2 + \frac{[bp'(p-1) - t]}{p bp' r} \sigma^2 \quad (7.4.6)
\end{aligned}$$

$$\begin{aligned}
& \text{Cov}(y_{ij.} - y_{i..})(y_{i..} - y_{...}) = \text{Cov}(y_{ij.} - y_{i..})(x_{i..} - x_{...}) = 0 \\
& \quad (7.4.7)
\end{aligned}$$

$$\begin{aligned}
& \text{Cov}(y_{i..} - y_{...})(x_{i..} - x_{...}) = E(y_{i..})(x_{i..}) - E(y_{i..})(x_{...}) \\
& \quad - E(y_{...})(x_{i..}) + E(y_{...})(x_{...}) = \frac{(b-1)}{br} \sigma_b^2 \quad (7.4.8)
\end{aligned}$$

The true deviations of the entries are the p 's defined in Model 7.1.3 which leads to the expectations:

$$E p(y_{ij.}) = \sigma_p^2 \quad (7.4.9)$$

$$E p(y_{i..}) = \frac{\sigma_p^2}{p'} \quad (7.4.10)$$

$$E p(y_{...}) = \frac{\sigma_p^2}{bp'} \quad (7.4.11)$$

where σ_p^2 is the covariance between the true values and the observed values of entry means and is equivalent to the component of variance for the entry effect, p_{ij} , defined in Model 7.1.3.

In terms of the expectations, 7.4.9 to 7.4.11, the covariances involved are

$$\text{Cov } p(y_{ij.} - y_{i..}) = E[p(y_{ij.})] - E[p(y_{i..})] = \frac{(p'-1)}{p'} \sigma_p^2 \quad (7.4.12)$$

$$\text{Cov } p(Y_{i..} - y_{...}) = E[p(y_{i..})] - E[p(y_{...})] = \frac{(b-1)}{bp'} \sigma_p^2 \quad (7.4.13)$$

$$\text{Cov } p(y_{ij.} - y_{...}) = E[p(y_{ij.})] - E[p(y_{...})] = \frac{(bp'-1)}{bp'} \sigma_p^2 \quad (7.4.14)$$

The true values of entry means are unrelated to the check deviations. Therefore,

$$\begin{aligned} \text{Cov } p[(y_{ij.} - y_{i..}) + \frac{t}{p}(y_{i..} - y_{...}) - \frac{t}{p}(x_{i..} - x_{...})] \\ = \text{Cov } p[(y_{ij.} - y_{i..}) + \frac{t}{p}(y_{i..} - y_{...})] = \frac{[bp'(p-1) - t]}{pbp'} \sigma_p^2 \end{aligned} \quad (7.4.15)$$

$$\text{Cov } p[(y_{ij.} - y_{i..}) - (x_{i..} - x_{...})] = \text{Cov } p(y_{ij.} - y_{i..}) \quad (7.4.16)$$

$$\text{Cov } p[(y_{ij.} - y_{...}) - (x_{i..} - x_{...})] = \text{Cov } p(y_{ij.} - y_{...}) \quad (7.4.17)$$

$$\text{Cov } p(x_{i..} - x_{...}) = 0 \quad (7.4.18)$$

If selection is based on entry means adjusted for block deviations of both checks and entries as outlined in Section 7.2, the expected gain is

$$\begin{aligned} \Delta_{lc} &= \frac{S_c \text{Cov } p[(y_{ij.} - y_{i..}) + \frac{t}{p}(y_{i..} - y_{...}) - \frac{t}{p}(x_{i..} - x_{...})]}{\sqrt{V[(y_{ij.} - y_{i..}) + \frac{t}{p}(y_{i..} - y_{...}) - \frac{t}{p}(x_{i..} - x_{...})]}} \\ &= \sqrt{\frac{T-b}{T} - \frac{(t/p)}{T(1-t/p)}} \frac{S_c \sigma_p^2}{\sqrt{[1 - \frac{t(T-p)(1-t/p)}{pT(p-1)(1-t/p) - t/T}] \sigma_p^2 + \frac{\sigma_r^2}{r}}} \end{aligned}$$

where S_c is the selection differential expressed in standard deviations.

For the special case where $t=0$, the expected gain from within blocks selection becomes:

$$\Delta_1 = \frac{S \text{Cov } p(y_{ij.} - y_{i..})}{\sqrt{V(y_{ij.} - y_{i..})}} = \sqrt{\frac{T-b}{T}} \frac{S \sigma_p^2}{\sqrt{\sigma_p^2 + \frac{\sigma^2}{r}}}$$

and if blocking is ignored, the expected gain from selection is

$$\begin{aligned} \Delta_2 &= \frac{S \text{Cov } p(y_{ij.} - y_{...})}{\sqrt{V(y_{ij.} - y_{...})}} = \sqrt{\frac{bp-1}{bp}} \frac{S \sigma_p^2}{\sqrt{\frac{p(b-1)}{r(bp-1)} \sigma_b^2 + \sigma_p^2 + \frac{\sigma^2}{r}}} \\ &= \sqrt{\frac{T-1}{T}} \frac{S \sigma_p^2}{\sqrt{\frac{\sigma_T^2}{r} + \sigma_p^2}} \end{aligned}$$

It should be remembered that $S_c \neq S$ if a constant number of top entries is selected, since the inclusion of check plots in an experiment necessarily eliminates some of the entries.

For selection on entry means adjusted for check deviations, the expected gain is

$$\begin{aligned} \Delta_{2c} &= \frac{S_c \text{Cov } p[(y_{ij.} - y_{...}) - (x_{i..} - x_{...})]}{\sqrt{V[(y_{ij.} - y_{...}) - (x_{i..} - x_{...})]}} \\ &= \sqrt{\frac{bp'-1}{bp'}} \frac{S_c \sigma_p^2}{\sqrt{\sigma_p^2 + \frac{[p'(bt+b-1)-t]}{\text{tr}(bp'-1)} \sigma^2}} \end{aligned}$$

The environmental components of variance in Δ_{1c} , Δ_1 , Δ_2 , and Δ_{2c} may be written in terms of the average error variances developed in

Section 7.3. That is,

$$\Delta_{1c} = \frac{b(p-1)(p-t)-t}{\sqrt{pb(p-t)}} \sqrt{\frac{s_c \sigma_p^2}{\frac{p(bp-bt-1)\sigma_{1c}^2}{2} + \left[\frac{p^2 \{b[p^2(p-t-1)+t^2]-t^2\}}{p^3} - \frac{(b-1)(p-t)^3(p-1)}{p^3} \right] \sigma_p^2}}$$

$$\Delta_1 = \sqrt{\frac{T-b}{T}} \sqrt{\frac{s \sigma_p^2}{\frac{(T-1)\sigma_1^2}{(T-b)2} + \frac{(T-b+1)}{T} \sigma_p^2}}$$

$$\Delta_2 = \sqrt{\frac{T-1}{T}} \sqrt{\frac{s \sigma_p^2}{\frac{\sigma_2^2}{2} + \sigma_p^2}}$$

$$\Delta_{2c} = \sqrt{\frac{bp'-1}{bp'}} \sqrt{\frac{s_c \sigma_p^2}{\frac{\sigma_{2c}^2}{2} + \sigma_p^2}}$$

The information previously formulated in equations 7.4.1 to 7.4.18 may be used to solve for the $\hat{\beta}$ values needed in the best index method of selection. In terms of variances and covariances, the normal equations to be solved are

$$\frac{p'-1}{p'r} (\sigma^2 + r\sigma_p^2) \beta_{1c} + (0) \beta_{2c} + (0) \beta_{3c} = \frac{(p'-1)}{p'} \sigma_p^2$$

$$(0) \beta_{1c} + \frac{b-1}{bp'r} (p'\sigma_b^2 + r\sigma_p^2 + \sigma^2) \beta_{2c} + \left[\frac{(b-1)}{br} \sigma_b^2 \right] \beta_{3c} = \frac{(b-1)}{bp'} \sigma_p^2$$

$$(0) \beta_{1c} + \left[\frac{(b-1)}{br} \sigma_b^2 \right] \beta_{2c} + \frac{b-1}{brt} (\sigma^2 + t\sigma_b^2) \beta_{3c} = 0$$

and the β 's are

$$\beta_{1c} = \frac{\sigma_p^2}{\sigma_p^2 + \frac{\sigma^2}{r}}$$

$$\beta_{2c} = \frac{r(\sigma^2 + t\sigma_b^2) \sigma_p^2}{(\sigma^2 + t\sigma_b^2)(\sigma^2 + p'\sigma_b^2 + r\sigma_p^2) - tp'(\sigma_b^2)^2}$$

$$\beta_{3c} = \frac{-tr(\sigma_b^2) \sigma_p^2}{(\sigma^2 + t\sigma_b^2)(\sigma^2 + p'\sigma_b^2 + r\sigma_p^2) - tp'(\sigma_b^2)^2}$$

The expected gain from selection for the index employing these values is

$$\begin{aligned} \Delta_{1c} &= S_c \sqrt{\beta_{1c} \text{Cov } p(y_{1j} - y_{1..}) + \beta_{2c} \text{Cov } p(y_{1..} - y_{...}) + \beta_{3c} \text{Cov } p(x_{1..} - x_{...})} \\ &= S_c \sigma_p^2 \sqrt{\frac{r(p'-1)}{p'(\sigma^2 + r\sigma_p^2)} + \frac{r(b-1)(\sigma^2 + t\sigma_b^2)}{bp'[(\sigma^2 + t\sigma_b^2)(\sigma^2 + p'\sigma_b^2 + r\sigma_p^2) - tp'(\sigma_b^2)^2]}} \end{aligned}$$

When $t=0$ the normal equations reduce to

$$\frac{p-1}{pr} (\sigma^2 + r\sigma_p^2) \beta_1 + (0) \beta_2 = \frac{(p-1)}{p} \sigma_p^2$$

$$(0) \beta_1 + \frac{b-1}{b} \left(\frac{\sigma_b^2}{r} + \frac{\sigma_p^2}{p} + \frac{\sigma^2}{pr} \right) \beta_2 = \frac{(b-1)}{bp} \sigma_p^2$$

and the β 's are

$$\beta_1 = \frac{\frac{\sigma_p^2}{p}}{\frac{\sigma_p^2}{p} + \frac{\sigma^2}{r}} = \beta_{1c}$$

$$\beta_2 = \frac{\frac{\sigma_p^2}{p}}{p \left(\frac{\sigma_b^2}{r} + \frac{\sigma_p^2}{p} + \frac{\sigma^2}{pr} \right)}$$

The expected gain for the best index method of selection when checks are not included is

$$\begin{aligned} \Delta_I &= S \sqrt{\beta_1 \text{Cov } p(y_{ij} - y_{i..}) + \beta_2 \text{Cov } p(y_{i..} - y_{...})} \\ &= S \sigma_p^2 \sqrt{\frac{r(p-1)}{p(\sigma^2 + r\sigma_p^2)} + \frac{r(b-1)}{bp(\sigma^2 + r\sigma_p^2 + p\sigma_b^2)}} \end{aligned}$$

7.5 Comparisons of Methods of Selection

A measure of the relative efficiencies, in terms of expected gain, for the alternative methods of selection may be obtained by expressing Δ_1 , Δ_I , Δ_{2c} , and Δ_{1c} as fractions of Δ_2 . Since Δ_2 is independent of block size, the magnitude of these ratios will be a direct reflection of the effect of block size on the other Δ 's.

The ratio of expected gain from within blocks selection to the expected gain from selection ignoring blocking is

$$\frac{\Delta_1}{\Delta_2} = \sqrt{\frac{T-b}{T-1}} \sqrt{\frac{r\sigma_p^2 + \sigma_T^2}{r\sigma_p^2 + \sigma^2}}.$$

For $1 < b < T$, $\sigma_T^2 > \sigma^2$, and since σ^2 becomes smaller as b increases, it follows that σ_T^2/σ^2 is a monotone increasing function. The fact that $r\sigma_p^2$ appears in both the numerator and denominator of the ratio in no way affects the monotonicity of this function, but only affects its relative magnitude.

The quantity $\sqrt{\frac{T-b}{T-1}} < 1$ for $1 < b < T$ and is a monotone decreasing function. The value of b which maximizes the ratio Δ_1/Δ_2 cannot be determined unless a specific relationship between block size and σ^2 is assumed.

In terms of average error variances the ratio of Δ_1 to Δ_2 is

$$\frac{\Delta_1}{\Delta_2} = \sqrt{\frac{T-b}{T-1}} \sqrt{\frac{\sigma_2^2 + 2\sigma_p^2}{\frac{(T-1)}{(T-b)}\sigma_1^2 + \frac{2(T-b+1)}{T}\sigma_p^2}}.$$

Thus, the ratio of Δ_1 to Δ_2 is not a monotonic function of the average error variances and the value of b that maximizes the ratio of these variances does not necessarily maximize the ratio of expected gains.

The ratio of expected gain from selection on entry means adjusted for block deviations of both entries and checks to the expected gain from selection ignoring blocks is

$$\frac{\Delta_{1c}}{\Delta_2} = \frac{S_c}{S} \sqrt{\frac{T-b}{T-1} - \frac{(t/p)}{(T-1)(1-t/p)}} \sqrt{\frac{r\sigma_p^2 + \sigma_T^2}{r[1 - \frac{t(T-p)(1-t/p)}{pT(p-1)(1-t/p)-t/T}]\sigma_p^2 + \sigma^2}}.$$

Except when t/p is close to one, the quantity in the left hand radical is approximately equal to $\frac{T-b}{T-1}$ and the coefficient of σ_p^2 in the denominator of the right hand radical is approximately equal to one.

Thus,

$$\frac{\Delta_{1c}}{\Delta_2} \doteq \frac{S_c}{S} \sqrt{\frac{T-b}{T-1}} \sqrt{\frac{r\sigma_p^2 + \sigma_T^2}{r\sigma_p^2 + \sigma^2}} = \frac{\Delta_1}{\Delta_2} \frac{S_c}{S}.$$

As in the previous comparisons, Δ_{1c}/Δ_2 is not a monotonic function of the average error variances.

A somewhat different result is obtained when Δ_{2c} , the expected gain from selection on entry means adjusted for check deviations, is compared to Δ_2 . This ratio is

$$\frac{\Delta_{2c}}{\Delta_2} = \frac{S_c}{S} \sqrt{\frac{T(T-bt-1)}{T(T-bt-1)+bt}} \sqrt{\frac{\sigma_T^2 + r\sigma_p^2}{\frac{[(t+1)(T-bt)-p]}{t(T-bt-1)} \sigma^2 + r\sigma_p^2}}.$$

In terms of average error variances the ratio is

$$\frac{\Delta_{2c}}{\Delta_2} = \frac{S_c}{S} \sqrt{1 - \frac{bt}{(T-1)(T-bt)}} \sqrt{\frac{\frac{\sigma_2^2}{2} + 2\sigma_p^2}{\frac{\sigma_{2c}^2}{2} + 2\sigma_p^2}}.$$

Since an experiment with blocks of less than two plots containing only check plots is unrealistic, in general $bt \leq T/2$ and the quantity in the left hand radical is approximately equal to one. Therefore,

$$\frac{\Delta_{2c}}{\Delta_2} \doteq \frac{S_c}{S} \sqrt{\frac{\frac{\sigma_2^2}{2} + 2\sigma_p^2}{\frac{\sigma_{2c}^2}{2} + 2\sigma_p^2}}.$$

That is, the ratio of Δ_{2c} to Δ_2 is a direct reflection of the ratio of the average error variances, except for differences in the selection differential.

The ratio of expected gain for the best index to the expected gain from selection ignoring blocks is:

$$\frac{\Delta_1}{\Delta_2} = \frac{1}{\sqrt{T-1}} \sqrt{\frac{(T-b)(\sigma_T^2 + r\sigma_p^2)}{\sigma^2 + r\sigma_p^2} + \frac{(b-1)(\sigma_T^2 + r\sigma_p^2)}{\sigma^2 + r\sigma_p^2 + p\sigma_b^2}}.$$

The value of b which will maximize this ratio depends on the relationship between block size and the environmental components of variance as in the previous comparisons. A useful approximation to this relationship is available in the expressions derived from Smith's variance law (7.1.5 and 7.1.6). If some arbitrary value is assigned to the ratio σ_p^2/σ_T^2 , then numerical values may be calculated for the Δ 's for various values of b , B , T , and r .

Values for Δ_1/Δ_2 are presented in Table 5 for the case where σ_p^2/σ_T^2 tends to zero, with $T=256$ and $r=2$. Table 5 indicates that the expected gain, Δ_1 , from within blocks selection is decidedly better than the expected gain, Δ_2 , when blocking is ignored as long as the correlation between adjacent plots within a block is high (B is small). With B values in the neighborhood of 0.5, the optimum number of blocks is in the range 16 to 32. As B becomes larger, within blocks selection becomes less and less efficient for all block sizes.

In Table 6, σ_p^2/σ_T^2 has been set equal to one-half. The result is a reduction in the ratio of Δ_1 to Δ_2 for all block sizes, except

Table 5. Ratio of the expected gain from within blocks selection to the expected gain from selection ignoring blocks, Δ_1/Δ_2 , for $T = 256$, $r = 2$, and σ_p^2/σ_T^2 approaching zero.

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	1.065	1.048	1.024	1.010	1.003	0.999	0.998
4	1.141	1.105	1.054	1.022	1.006	0.997	0.992
8	1.230	1.172	1.089	1.038	1.008	0.992	0.986
16	1.331	1.247	1.127	1.052	1.006	0.979	0.968
32	1.434	1.323	1.161	1.058	0.993	0.952	0.937
64	1.506	1.366	1.162	1.031	0.946	0.889	0.866
128	1.420	1.267	1.042	0.898	0.801	0.734	0.708

Table 6. Ratio of the expected gain from within blocks selection to the expected gain from selection ignoring blocks, Δ_1/Δ_2 , for $T = 256$, $r = 2$, and $\sigma_p^2/\sigma_T^2 = 0.5$.

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	1.030	1.022	1.011	1.004	1.000	0.998	0.998
4	1.060	1.045	1.023	1.008	1.000	0.995	0.992
8	1.088	1.067	1.034	1.011	0.997	0.989	0.986
16	1.109	1.083	1.040	1.009	0.988	0.975	0.968
32	1.110	1.082	1.031	0.992	0.964	0.944	0.937
64	1.063	1.036	0.983	0.939	0.904	0.878	0.866
128	0.896	0.874	0.829	0.787	0.750	0.721	0.708

when B is equal to one. For B values near 0.5, the optimum block number is in the range 8 to 16. Thus, any increase in the relative size of the entry component of variance tends to reduce the ratio of Δ_1 to Δ_2 , and increases the optimum block size.

The exact points at which $\Delta_1 = \Delta_2$ may be determined by solving the equation $\Delta_1^2 - \Delta_2^2$ for either B or b.

$$\begin{aligned}\Delta_1^2 - \Delta_2^2 &= \frac{T-b}{T(\sigma_p^2 + \frac{\sigma^2}{r})} - \frac{T-1}{T(\sigma_p^2 + \frac{\sigma_T^2}{r})} \\ &= b^2[T^B - 1 + k(T^B - 1)] + b[2(T - T^{B+1}) + k(T - T^B - T^{B+1} + 1)] \\ &\quad + b^B(1 - 2T + T^2) - T^2 + 2T^{B+1} - T^B + k(T^{B+1} - T) = 0\end{aligned}$$

where $k = r\sigma_p^2/\sigma_T^2$. The solution is found most easily by substituting in various values of B and then solving for b by methods of numerical analysis.

The values of b and B for which $\Delta_1 > \Delta_2$, $\Delta_1 = \Delta_2$, and $\Delta_1 < \Delta_2$ are shown in the graph in Figure 1 for $T = 256$, $r = 2$ and σ_p^2/σ_T^2 approaching zero. Figure 2 gives the analagous situation when $\sigma_p^2/\sigma_T^2 = 0.5$.

The ratios of Δ_1 to Δ_2 for $T = 256$, $r = 2$, and σ_p^2/σ_T^2 approaching zero are presented in Table 7. The general response is much the same as that in Table 5, but in this case the ratios never become less than one. In addition, the optimum block number remains constant at 64 for all values of B. This is in contrast to the Δ_1/Δ_2 comparison in which optimum block number decreased as B increased.

Figure 1. Graph of the equation $\Delta_1 - \Delta_2 = 0$ for $T = 256$, $r = 2$ and σ_p^2/σ_T^2 approaching zero

Figure 2. Graph of the equation $\Delta_1 - \Delta_2 = 0$ for $T = 256$, $r = 2$, and $\sigma_p^2/\sigma_T^2 = 0.5$

If σ_p^2/σ_T^2 , is set equal to 0.5, the ratios of Δ_1 to Δ_2 are decreased as shown in Table 8. The optimum block number is in the neighborhood of 32.

When check information is included in the index, the expected gain from selection is approximately equal to the expected gain for an index without checks, except for differences in the selection differential.

That is,

$$\frac{\Delta_{1c}}{\Delta_1} \approx \frac{S_c}{S}.$$

Values for Δ_{2c}/Δ_2 are presented in Table 9 for $T = 256$, $r = 2$, and σ_p^2/σ_T^2 , approaching zero, when the top 25 entries are selected. Table 10 gives values for the same experiment when $\sigma_p^2/\sigma_T^2 = 0.5$.

In each case the expected gain for entry means adjusted for checks is less than the expected gain when checks are not included, except when B is very small. Since values of $B \leq 0.1$ are probably unrealistic under ordinary field conditions, the inclusion of checks for the purpose of adjusting entry means does not appear to be a desirable alternative.

7.6 Effect of Replication of Entries on Expected Gain from Selection

In the examples discussed in the preceding sections, the number of replications involved has been held constant and only variable block size has been considered. However, it is also of interest to consider the effect of replication. If an experiment consists of a fixed number of plots, the inclusion of additional replications must necessarily

Table 7. Ratio of the expected gain from selection using the best index to the expected gain from selection ignoring blocks, Δ_1/Δ_2 , for $T = 256$, $r = 2$, and σ_p^2/σ_T^2 , approaching zero

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	1.065	1.048	1.024	1.010	1.003	1.000	1.000
4	1.141	1.106	1.054	1.024	1.008	1.001	1.000
8	1.231	1.173	1.090	1.041	1.013	1.002	1.000
16	1.334	1.250	1.132	1.060	1.020	1.002	1.000
32	1.442	1.333	1.175	1.080	1.026	1.003	1.000
64	1.533	1.397	1.206	1.092	1.030	1.003	1.000
128	1.516	1.378	1.187	1.080	1.025	1.002	1.000

Table 8. Ratio of the expected gain from selection using the best index to the expected gain from selection ignoring blocks, Δ_1/Δ_2 , for $T = 256$, $r = 2$, and $\sigma_p^2/\sigma_T^2 = 0.5$

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	1.030	1.022	1.011	1.004	1.001	1.000	1.000
4	1.060	1.046	1.024	1.010	1.003	1.000	1.000
8	1.090	1.069	1.037	1.016	1.008	1.000	1.000
16	1.114	1.089	1.049	1.022	1.006	1.000	1.000
32	1.127	1.102	1.057	1.026	1.008	1.001	1.000
64	1.119	1.096	1.056	1.026	1.004	1.001	1.000
128	1.079	1.065	1.038	1.018	1.006	1.000	1.000

Table 9. Ratio of the expected gain from selection among entry means adjusted for check deviations to the expected gain from selection without checks, Δ_{2c}/Δ_2 , for $T = 256$, $r = 2$, and σ_p^2/σ_T^2 approaching zero; when the top 25 entries are selected

For $t = 2$

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	0.954	0.939	0.917	0.905	0.898	0.895	0.894
4	0.978	0.947	0.903	0.877	0.862	0.854	0.852
8	1.040	0.990	0.920	0.877	0.852	0.838	0.833
16	1.102	1.034	0.934	0.871	0.833	0.811	0.803
32	1.165	1.074	0.943	0.859	0.806	0.773	0.761
64	1.129	1.025	0.871	0.773	0.709	0.666	0.651

For $t = 4$

2	1.006	0.990	0.967	0.954	0.947	0.943	0.942
4	1.053	1.020	0.972	0.944	0.928	0.920	0.917
8	1.100	1.048	0.974	0.928	0.902	0.887	0.882
16	1.146	1.074	0.970	0.906	0.866	0.843	0.835
32	1.092	1.007	0.884	0.805	0.756	0.724	0.714

Table 10. Ratio of the expected gain from selection among entry means adjusted for check deviations to the expected gain from selection without checks, Δ_{2c}/Δ_2 , for $T = 256$, $r = 2$, and $\sigma_p^2/\sigma_T^2 = 0.5$; when the top 25 entries are selected

For $t = 2$

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	0.977	0.969	0.957	0.949	0.945	0.943	0.942
4	0.990	0.974	0.949	0.933	0.924	0.919	0.917
8	1.018	0.995	0.958	0.933	0.917	0.908	0.905
16	1.028	1.001	0.954	0.919	0.896	0.881	0.876
32	1.020	0.990	0.935	0.892	0.861	0.840	0.832
64	0.911	0.884	0.830	0.784	0.749	0.723	0.712

For $t = 4$

2	1.003	0.995	0.984	0.976	0.972	0.970	0.970
4	1.024	1.009	0.986	0.971	0.962	0.957	0.956
8	1.028	1.008	0.974	0.950	0.936	0.927	0.925
16	1.014	0.990	0.948	0.917	0.896	0.882	0.878
32	0.902	0.879	0.835	0.800	0.775	0.758	0.751

result in a smaller number of entries tested. Whether or not this is desirable can best be determined by considering the effect of this variable on the expected gain from selection. It was pointed out in Section 7.4 that varying the number of entries tested results in a change in the selection differential when a constant number of top entries is selected. The magnitude of this change is also dependent upon the intensity of the selection pressure.

The ratio Δ_1/Δ_1 is presented in the graph in Figure 3 for an experiment of 512 plots, where Δ_1 is the expected gain from selection within blocks using 2 replications, and Δ_1 is the expected gain using 4 replications. The values in the graph are for a block of 16 plots with $B=0.5$. It should be emphasized, however, that for the BIR design the relative effect on expected gain of varying number of replications is only very slightly dependent upon block size and the size of the regression coefficient, B . This is particularly true for B values in the range 0.3 to 0.7.

The graph indicates that additional replication is advantageous ($\Delta_1/\Delta_1 < 1$) only with more intensive selection and when σ_p^2 is very small. Graphs for Δ_2/Δ_2 and Δ_1/Δ_1 are almost identical to Figure 3 and are, therefore, not included.

The effect of replication on the average error variance of pairwise entry comparisons is in sharp contrast to this result, since the use of additional replications always results in a decrease in average error variance.

Figure 3. Graph of the ratio Δ_1/Δ_1 for an experiment of 512 plots with $b=16$ and $B=0.5$, where Δ_1 is the expected gain from selection within blocks using 2 replications and Δ_1 is the expected gain using 4 replications.

7.7 Inclusion of Family Structure in Selection Schemes Involving the BIR Design

New complications are introduced when the entries have a pattern of relationships such as the full and half sib families of the Design I of Comstock and Robinson (1948). The appropriate model becomes:

$$Y_{ijkl} = u + r_k + b_{ki} + s_{ij} + d_{ijl} + e_{ijkl} . \quad (7.7.1)$$

The effects are r for replications, b for blocks in replications, s for males in blocks, d for females in males in blocks, and e for plots. The subscript i varies from one to b , j varies from one to s , k varies from one to r , and l varies from one to d . All effects are assumed to be random and uncorrelated. If selection on full sib family means is practiced, then for an experiment without checks, the methods of selection outlined in Section 7.2 may be included in a general prediction equation as follows:

$$\hat{Y}_{ij.l} = y_{....} + \hat{\beta}_{1f}(y_{ij.l} - y_{ij..}) + \hat{\beta}_{2f}(y_{ij..} - y_{i1...}) + \hat{\beta}_{3f}(y_{i1...} - y_{....})$$

where Y_{ijkl} is the value of progeny from the l th female mated to the j th sire in i th block in the k th replication, and the $\hat{\beta}$'s are weights to be assigned to the appropriate deviations.

The first deviation is that of the full sib family mean from the half sib family mean, the second deviation that of the half sib family mean from the block mean, and the third deviation that of the block mean from the over-all mean.

If selection within blocks on full sib family means is practiced, the $\hat{\beta}$'s are

$$\hat{\beta}_{1f} = \hat{\beta}_{2f} = 1, \quad \hat{\beta}_{3f} = 0$$

and the predicted value of a full sib family mean is

$$\hat{y}_{ij.1} = y_{....} + (y_{ij.1} - y_{i....})$$

When blocks are ignored and selection is on full sib family means, the $\hat{\beta}$'s take the values:

$$\hat{\beta}_{1f} = \hat{\beta}_{2f} = \hat{\beta}_{3f} = 1$$

and

$$\hat{y}_{ij.1} = y_{ij.1}$$

The following expectations are used in deriving the variances and covariances of the deviations and are based on Model 7.7.1:

$$E(y_{ij.1})(y_{ij..}) = E(y_{ij..})^2 = u^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{r} + \sigma_s^2 + \frac{\sigma_d^2}{d} + \frac{\sigma^2}{rd}$$

$$\begin{aligned} E(y_{ij.1})(y_{i....}) &= E(y_{ij..})(y_{i....}) = E(y_{i....})^2 \\ &= u^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{r} + \frac{\sigma_s^2}{s} + \frac{\sigma_d^2}{sd} + \frac{\sigma^2}{rsd} \end{aligned}$$

$$\begin{aligned} E(y_{ij.1})(y_{....}) &= E(y_{ij..})(y_{....}) = E(y_{i....})(y_{....}) = E(y_{....})^2 \\ &= u^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{br} + \frac{\sigma_s^2}{bs} + \frac{\sigma_d^2}{bsd} + \frac{\sigma^2}{bsdr} \end{aligned}$$

$$E(y_{ij.1})^2 = u^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_b^2}{r} + \sigma_s^2 + \sigma_d^2 + \frac{\sigma^2}{r}$$

In terms of these expectations, the variances and covariances of the deviations are

$$V(y_{ij.l} - y_{ij..}) = E(y_{ij.l})^2 - E(y_{ij..})^2 = \frac{d-1}{d} (\sigma_d^2 + \frac{\sigma_r^2}{r})$$

$$V(y_{ij..} - y_{i....}) = E(y_{ij..})^2 - E(y_{i....})^2 = \frac{s-1}{s} (\sigma_s^2 + \frac{\sigma_d^2}{d} + \frac{\sigma_{rd}^2}{rd})$$

$$V(y_{i....} - y_{.....}) = E(y_{i....})^2 - E(y_{.....})^2 = \frac{b-1}{b} (\frac{\sigma_b^2}{r} + \frac{\sigma_s^2}{s} + \frac{\sigma_d^2}{sd} + \frac{\sigma_{rsd}^2}{rsd})$$

$$V(y_{ij.l} - y_{i....}) = E(y_{ij.l})^2 - E(y_{i....})^2 = \frac{(s-1)}{s} \sigma_s^2 + \frac{(sd-1)}{sd} \sigma_d^2 + \frac{(sd-1)}{rsd} \sigma_{rsd}^2$$

$$V(y_{ij.l} - y_{.....}) = E(y_{ij.l})^2 - E(y_{.....})^2 = \frac{(b-1)}{br} \sigma_b^2 + \frac{(bs-1)}{bs} \sigma_s^2 + \frac{(bsd-1)}{bsd} \sigma_d^2 + \frac{(bsd-1)}{bsdr} \sigma_{rsd}^2$$

$$\begin{aligned} \text{Cov}(y_{ij.l} - y_{ij..})(y_{ij..} - y_{i....}) &= \text{Cov}(y_{ij.l} - y_{ij..})(y_{i....} - y_{.....}) \\ &= \text{Cov}(y_{ij..} - y_{i....})(y_{i....} - y_{.....}) = 0. \end{aligned}$$

If the parents selected are to be recombined in the next generation, or if the genetic variance is all additive, the covariances needed to derive the expected gain from selection are the covariances between the deviations and the additive genetic values, G 's, of full sib families. When the parents selected are remnant seed, i.e., full sibs of those used in the test, the expectations involved in deriving these covariances are

$$E G(y_{ij.1}) = \frac{\sigma_G^2}{2}$$

$$E G(y_{ij..}) = \frac{\sigma_G^2}{2d} + \frac{(d-1)}{4d} \sigma_G^2 = \frac{(d+1)}{4d} \sigma_G^2$$

$$E G(y_{i...}) = \frac{\sigma_G^2}{2sd} + \frac{(d-1)}{4sd} \sigma_G^2 = \frac{(d+1)}{4sd} \sigma_G^2$$

$$E G(y_{....}) = \frac{\sigma_G^2}{2bsd} + \frac{(d-1)}{4bsd} \sigma_G^2 = \frac{(d+1)}{4bsd} \sigma_G^2$$

where σ_G^2 = additive genetic variance. The genetic covariances are

$$\text{Cov } G(y_{ij.1} - y_{ij..}) = E G(y_{ij.1}) - E G(y_{ij..}) = \frac{(d-1)}{4d} \sigma_G^2$$

$$\text{Cov } G(y_{ij..} - y_{i...}) = E G(y_{ij..}) - E G(y_{i...}) = \frac{(s-1)(d+1)}{4sd} \sigma_G^2$$

$$\text{Cov } G(y_{i...} - y_{....}) = E G(y_{i...}) - E G(y_{....}) = \frac{(b-1)(d+1)}{4bsd} \sigma_G^2$$

$$\text{Cov } G(y_{ij.1} - y_{i...}) = E G(y_{ij.1}) - E G(y_{i...}) = \frac{[2sd - (d+1)]}{4sd} \sigma_G^2$$

$$\text{Cov } G(y_{ij.1} - y_{....}) = E G(y_{ij.1}) - E G(y_{....}) = \frac{[2bsd - (d+1)]}{4bsd} \sigma_G^2$$

For selection in which the full sib family means are adjusted for block means, the expected gain is

$$\Delta_{1f} = \frac{s \sigma_G^2 [2sd - (d+1)]}{4sd \sqrt{\frac{(s-1)}{s} \sigma_s^2 + \frac{(sd-1)}{sd} \sigma_d^2 + \frac{(sd-1)}{rsd} \sigma_r^2}}$$

If blocks are ignored, the expected gain from selection on full family means is

$$\Delta_{2f} = \frac{s \sigma_G^2 [2bsd - (d+1)]}{4bsd \sqrt{\frac{(b-1)}{br} \sigma_b^2 + \frac{(bs-1)}{bs} \sigma_s^2 + \frac{(bsd-1)}{bsd} \sigma_d^2 + \frac{(bsd-1)}{bsdr} \sigma_r^2}}$$

The normal equations necessary to solve for the β 's in the best index are

$$\frac{d-1}{d} (\sigma_d^2 + \frac{\sigma_r^2}{r}) \beta_{1f} + (0) \beta_{2f} + (0) \beta_{3f} = \frac{(d-1)}{4d} \sigma_G^2$$

$$(0) \beta_{1f} + \frac{s-1}{s} (\sigma_s^2 + \frac{\sigma_d^2}{d} + \frac{\sigma_{rd}^2}{rd}) \beta_{2f} + (0) \beta_{3f} = \frac{(s-1)(d+1)}{4sd} \sigma_G^2$$

$$(0) \beta_{1f} + (0) \beta_{2f} + \frac{b-1}{b} (\frac{\sigma_b^2}{r} + \frac{\sigma_s^2}{s} + \frac{\sigma_d^2}{sd} + \frac{\sigma_{rsd}^2}{rsd}) \beta_{3f} = \frac{(b-1)(d+1)}{4bsd} \sigma_G^2 .$$

From these equations the β 's are found to be

$$\beta_{1f} = \frac{\sigma_G^2}{4(\sigma_d^2 + \frac{\sigma_r^2}{r})}$$

$$\beta_{2f} = \frac{(d+1) \sigma_G^2}{4d(\sigma_s^2 + \frac{\sigma_d^2}{d} + \frac{\sigma_{rd}^2}{rd})}$$

$$\beta_{3f} = \frac{(d+1) \sigma_G^2}{4sd(\frac{\sigma_b^2}{r} + \frac{\sigma_s^2}{s} + \frac{\sigma_d^2}{sd} + \frac{\sigma_{rsd}^2}{rsd})} .$$

The expected gain for the best index method of selection is

$$\Delta_{If} = s \sigma_G^2 \sqrt{\frac{d-1}{16d(\sigma_d^2 + \frac{\sigma_r^2}{r})} + \frac{(d+1)^2}{16bs^2d^2} \left[\frac{bs(s-1)}{\sigma_s^2 + \frac{\sigma_d^2}{d} + \frac{\sigma_{rd}^2}{rd}} + \frac{b-1}{\frac{\sigma_b^2}{r} + \frac{\sigma_s^2}{s} + \frac{\sigma_d^2}{sd} + \frac{\sigma_{rsd}^2}{rsd}} \right]} .$$

It should be noted that an additional variable enters into the expressions for expected gain when family structure is included in the design. This variable is d , the number of females mated to the same

male within a block. The permissible range of d is from one to p , where $p=sd$ is the block size.

To avoid the confounding effect of block number, it is useful to look at the expressions for expected gain for various values of d when an experiment consists of a single block. If the components of variance are expressed as fractions of the total variance, σ_T^2 , as in Section 7.1, it is possible to compute numerical values (in terms of units of expected gain) for expected gain so that graphical comparisons can be made.

Figure 4(a) is a graphical representation of the expected gain from selection (with d variable) for the alternative methods when an experiment consists of a single block of 64 progenies, with 2 replications. Note that when $b=1$, $\Delta_{1f} = \Delta_{2f}$. The male and female components are assumed to be of equal size, the situation when genetic variance is all additive, with $\sigma_d^2/\sigma_T^2 = \sigma_s^2/\sigma_T^2 = 0.1$, and $B = 0.5$.

The graph indicates that when $d=1$ and $d=64$, the extreme values of d , $\Delta_{1f} = \Delta_{2f} = \Delta_{1f}$ as would be expected. The maximum expected gain from selection on full sib family means is achieved when $d=1$, i.e., no half sib families are included. This is the unrelated progeny situation of Section 7.4. This result is expected, since the inclusion of half sib families in a block causes the comparisons within a block to be among related families.

The maximum expected gain for the best index is realized when 8 males with 8 females per male are used. With this allocation of males and females, the index method is 23% more efficient than selecting only on full sib family means.

The effect on expected gain when the experiment is subdivided into smaller blocks is indicated in Figure 4 (b,c,d,e,f). In each case, the expected gain from selection within blocks decreases rapidly as the number of half sib families increases. When blocks are ignored in selection, the effect is much less drastic.

Note that, for $b=1,2,4$, the maximum expected gain for the index method of selection is achieved when 8 females per male are used. For smaller blocks, the optimum number of females per male is somewhat less. Except for the extreme case when $b=32$ (2 plots per block), the expected gain for the index is enhanced by the inclusion of half sib families to an optimum after which the gain is decreased by increasing the ratio of half sibs to full sibs.

It should be pointed out that increasing the size of the experiment and/or the relative sizes of the male and female components of variance does not significantly change the patterns of expected gains exhibited in Figure 4.

Figure 4. Graphical representation of the expected gain from selection for $T = 64$, $r = 2$, $\sigma_s^2/\sigma_T^2 = \sigma_d^2/\sigma_T^2 = 0.1$, and $B = 0.5$, where Δ_{1f} = expected gain from selection within blocks on full sib family means, Δ_{2f} = expected gain from selection ignoring blocks, and Δ_{1f} = expected gain from selection using an index.

Figure 4 (continued)

8.0 THE REPLICATION IN BLOCKS (RIB) DESIGN

8.1 Model and Analysis of Variance

Since the replication in blocks design described for the Design I mating system of Comstock and Robinson (1948) is often used in genetic experiments, it is of interest to consider this design as to its usefulness in selection.

Consider the model:

$$Y_{ijk} = u + B_i + R_{ik} + p_{ij} + e_{ijk}. \quad (8.1.1)$$

The effects are B for blocks, R for replications in blocks, p for entries in blocks, and e for plots. The subscript i varies from one to b, j varies from one to p, and k varies from one to r. All effects are assumed to be random and uncorrelated which leads to the analysis of variance in Table 11.

Table 11. Analysis of variance for the RIB design with expectations of mean squares for Model 8.1.1 and Model 8.1.2

Source of variation	d.f.	Expectations of mean squares	
		Model 8.1.1	Model 8.1.2
Blocks	b-1	$\sigma^2 + r\sigma_p^2 + p\sigma_r^2,$ $+ p\sigma_b^2,$	$\sigma^2 + p\sigma_r^2 + pr\sigma_b^2,$
Reps in blocks	b(r-1)	$\sigma^2 + p\sigma_r^2,$	$\sigma^2 + p\sigma_r^2,$
Entries in blocks	b(p-1)	$\sigma^2 + r\sigma_p^2$	σ^2
Ent. x reps in bl.	b(p-1)(r-1)	σ^2	σ^2

If uniformity of entries is assumed, the model becomes:

$$Y_{ijk} = u + B_i + R_{ik} + e_{ijk} . \quad (8.1.2)$$

The effects are B for blocks, R for replications in blocks, and e for plots. The subscripts are the same as in Model 8.1.1. The expectations of mean squares for Model 8.1.2 are given in Table 11.

From the analysis of variance in Table 11, the total environmental variance within a block, $\sigma_{T''}^2$, may be expressed as a function of σ^2 and $\sigma_{r'}^2$:

$$\sigma_{T''}^2 = \sigma^2 + \frac{p(r-1)}{rp-1} \sigma_{r'}^2 .$$

Solving for $\sigma_{r'}^2$, results in

$$\sigma_{r'}^2 = \frac{rp-1}{p(r-1)} (\sigma_{T''}^2 - \sigma^2) .$$

But in terms of Smith's variance function (6.0.6):

$$\sigma_{T''}^2 = \frac{rp[1-(pr)^{-B}](v_1)_{\infty}}{(rp-1)}$$

$$\sigma^2 = \frac{p(1-p^{-B})(v_1)_{\infty}}{(p-1)} .$$

Similarly, the total variance, $\sigma_{T'}^2$, within an experiment of $T' = rpb$ plots is

$$\sigma_{T'}^2 = \frac{br(p-1) \sigma^2 + b(r-1)(\sigma^2 + p\sigma_{r'}^2) + (b-1)(\sigma^2 + p\sigma_{r'}^2 + pr\sigma_b^2)}{(bpr - 1)}$$

and

$$\sigma_{b'}^2 = \frac{(bpr-1) \sigma_{T'}^2 - (bpr-1) \sigma^2 - p(br-1) \sigma_{r'}^2}{pr(b-1)}$$

$$= \frac{(r-1)(brp-1) \sigma_{T'}^2 + r(b-1)(p-1) \sigma^2 - (br-1)(rp-1) \sigma_{T''}^2}{pr(b-1)(r-1)}$$

where

$$\sigma_{T'}^2 = \frac{T' [1 - (T')^{-B}] (v_1)_\infty}{T' - 1}.$$

As in the previous design, σ^2 and $\sigma_{b'}^2$ are again functions of block size. However, in addition, $\sigma_{r'}^2$ is also dependent upon block size.

When expressed as a fraction of the total variance, $\sigma_{T'}^2$, these components of variance are

$$\frac{\sigma^2}{\sigma_{T'}^2} = \frac{T' - 1}{r(T-b)} \left[\frac{(T')^B - (br)^B}{(T')^B - 1} \right] \quad (8.1.3)$$

$$\frac{\sigma_{r'}^2}{\sigma_{T'}^2} = \frac{T' - b}{T(r-1)} \left[\frac{\sigma_{T''}^2}{\sigma_{T'}^2} - \frac{\sigma^2}{\sigma_{T'}^2} \right] \quad (8.1.4)$$

$$\frac{\sigma_{b'}^2}{\sigma_{T'}^2} = \frac{b(r-1)(T'-1) + r(b-1)(T-b) \frac{\sigma^2}{\sigma_{T'}^2} - (br-1)(T'-b) \frac{\sigma_{T''}^2}{\sigma_{T'}^2}}{Tr(b-1)(r-1)} \quad (8.1.5)$$

where

$$\frac{\sigma_{T''}^2}{\sigma_{T'}^2} = \frac{T' - 1}{T' - b} \left[\frac{(T')^B - b^B}{(T')^B - 1} \right] \quad (8.1.6)$$

and

$$\lim_{B \rightarrow \infty} \frac{\sigma^2}{\sigma_{T'}^2} = \frac{T'-1}{r(T-b)} \left[1 - \frac{\log_e (br)}{\log_e (T')} \right]$$

$$\lim_{B \rightarrow \infty} \frac{\sigma_{T''}^2}{\sigma_{T'}^2} = \frac{T'-1}{T'-b} \left[1 - \frac{\log_e b}{\log_e T'} \right].$$

8.2 Average Error Variance of Entry Comparisons

Unlike the BIR design, the average error variance per comparison of unadjusted entry means is not independent of block size.

The total number of pair-wise entry comparisons within and between blocks and the variance of each of these comparisons is given in Table 12.

Table 12. Total number of pair-wise entry comparisons between and within blocks of the RIB design and the variance of each of these comparisons

Type of comparison	Total no.	Variance of the comparison
Within block	$\frac{bp(p-1)}{2}$	$\frac{2\sigma^2}{r}$
Between blocks	$\frac{b(b-1)p^2}{2}$	$\frac{2(\sigma^2 + \sigma_{r'}^2 + r\sigma_{b'}^2)}{r}$

The average error variance per comparison is

$$\sigma_{2'}^2 = \frac{\frac{bp(p-1)}{r} \sigma^2 + \frac{b(b-1)p^2}{r} (\sigma^2 + \sigma_{r'}^2 + r\sigma_{b'}^2)}{\frac{bp(bp-1)}{2}} = \frac{2}{r} \left[\sigma^2 + \frac{p(b-1)}{bp-1} (\sigma_{r'}^2 + r\sigma_{b'}^2) \right].$$

If σ_r^2 , and σ_b^2 , are written as functions of σ^2 , $\sigma_{T'}^2$, and $\sigma_{T''}^2$, then

$$\begin{aligned}\bar{\sigma}_2^2 &= \frac{2}{r} \left\{ \sigma^2 + \frac{p(b-1)}{bp-1} \left[\frac{rp-1}{p(r-1)} (\sigma_{T''}^2 - \sigma^2) \right. \right. \\ &\quad \left. \left. + \frac{(r-1)(brp-1)\sigma_{T'}^2 + r(b-1)(p-1)\sigma^2 - (br-1)(rp-1)\sigma_{T''}^2}{p(b-1)(r-1)} \right] \right\} \\ &= \frac{2}{r(bp-1)} [b(p-1)\sigma^2 + b(1-rp)\sigma_{T''}^2 + (brp-1)\sigma_{T'}^2] .\end{aligned}$$

To determine the effect of block size on this expression, it is useful to write $\bar{\sigma}_2^2$ as a fraction of the total variance, $\sigma_{T'}^2$:

$$\frac{\bar{\sigma}_2^2}{\sigma_{T'}^2} = \frac{2}{r(bp-1)} [b(p-1) \frac{\sigma^2}{\sigma_{T'}^2} + b(1-rp) \frac{\sigma_{T''}^2}{\sigma_{T'}^2} + (brp-1)] .$$

If the variance components are expressed in terms of T' , b , r , and B as in equations 8.1.3 and 8.1.6, then

$$\frac{\bar{\sigma}_2^2}{\sigma_{T'}^2} = \frac{2(T'-1)}{r(T-1)} \left\{ \frac{1}{r} \left[\frac{(T')^B - (br)^B}{(T')^B - 1} \right] - \left[\frac{(T')^B - b^B}{(T')^B - 1} \right] + 1 \right\} .$$

Note that when the regression coefficient, B , is equal to one, $\bar{\sigma}_2^2/\sigma_{T'}^2$ is a constant. However, for $0 < B < 1$, the average error variance per comparison is dependent upon block size. When $r \geq 2$, the quantity

$$\frac{1}{r} \left[\frac{(T')^B - (br)^B}{(T')^B - 1} \right] - \left[\frac{(T')^B - b^B}{(T')^B - 1} \right]$$

is always negative and increases in size as block size increases.

Therefore, the average error variance per comparison decreases as block

size increases and is a minimum when $b=1$. This is the randomized complete block situation.

The rate at which $\bar{\sigma}_2^2$, changes with variable block size is illustrated in Table 13 for $T=256$ and $r=2$. The effect of block size on $\bar{\sigma}_2^2$, is most pronounced when the correlation between adjacent plots within a block is high (B is small).

The average error variances per comparison for the other alternative methods of selection are the same as for the BIR design, that is

$$\bar{\sigma}_{1'}^2 = \bar{\sigma}_1^2$$

$$\bar{\sigma}_{1c'}^2 = \bar{\sigma}_{1c}^2$$

$$\bar{\sigma}_{2c'}^2 = \bar{\sigma}_{2c}^2$$

Table 13. Ratio of the average error variance per comparison of unadjusted treatment means to the total variance, $\bar{\sigma}_2^2/\sigma_T^2$, for $T=256$ and $r=2$

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
1	0.891	0.919	0.960	0.983	0.994	0.999	1.000
2	1.002	0.997	0.992	0.994	0.997	0.999	1.000
4	1.113	1.078	1.032	1.010	1.002	1.000	1.000
8	1.225	1.167	1.081	1.032	1.010	1.001	1.000
16	1.336	1.261	1.142	1.064	1.023	1.004	1.000
32	1.447	1.364	1.216	1.109	1.044	1.009	1.000
64	1.559	1.473	1.308	1.173	1.078	1.019	1.000
128	1.670	1.590	1.420	1.263	1.133	1.037	1.000

8.3 Expected Gain from Selection

The same alternative methods of selection may be compared for the RIB design as for the BIR design by utilizing the prediction equation developed in Section 7.2, where Y_{ijk} is now the value of the j th entry in the k th replication in the i th incomplete block and X_{ilk} is the value of the l th check in the k th replication in the i th incomplete block. The following expectations are utilized in deriving the variances and covariances of the deviations when checks are not included:

$$\begin{aligned} E(Y_{ij.})(y_{...}) &= E(y_{i..})(y_{...}) = E(y_{...})^2 \\ &= u^2 + \frac{\sigma_b^2}{b} + \frac{\sigma_r^2}{br} + \frac{\sigma_p^2}{bp} + \frac{\sigma^2}{bpr} \\ E(Y_{ij.})(y_{i..}) &= E(y_{i..})^2 = u^2 + \sigma_b^2 + \frac{\sigma_r^2}{r} + \frac{\sigma_p^2}{p} + \frac{\sigma^2}{pr} \\ E(y_{ij.})^2 &= u^2 + \sigma_b^2 + \frac{\sigma_r^2}{r} + \sigma_p^2 + \frac{\sigma^2}{r} . \end{aligned}$$

The variances and covariances of the deviations are

$$\begin{aligned} V(y_{ij.}-y_{i..}) &= E(y_{ij.})^2 - E(y_{i..})^2 = \frac{p-1}{pr} (\sigma^2 + r\sigma_p^2) \\ V(y_{i..}-y_{...}) &= E(y_{i..})^2 - E(y_{...})^2 = \frac{b-1}{bpr} (\sigma^2 + r\sigma_p^2 + p\sigma_r^2 + pr\sigma_b^2) \\ V(y_{ij.}-y_{...}) &= E(y_{ij.})^2 - E(y_{...})^2 = \frac{(b-1)}{b} (\sigma_b^2 + \frac{\sigma_r^2}{r}) + \frac{bp-1}{bpr} (\sigma^2 + r\sigma_p^2) \\ \text{Cov}(y_{ij.}-y_{i..})(y_{i..}-y_{...}) &= E(y_{ij.})(y_{i..}) - E(y_{ij.})(y_{...}) \\ &\quad - E(y_{i..})^2 + E(y_{i..})(y_{...}) = 0 . \end{aligned}$$

The covariances between the true values and predicted values of entry means are the same as those derived for the BIR design (equations 7.4.12 to 7.4.14).

If blocking is ignored, the expected gain from selection is

$$\Delta_{2'} = S \sqrt{\frac{T-1}{T}} \frac{\sigma_p^2}{\sqrt{\frac{T(b-1)}{b(T-1)} (\sigma_{b'}^2 + \frac{\sigma_{r'}^2}{r}) + \frac{\sigma_r^2}{r} + \sigma_p^2}} .$$

In terms of average error variance this expression is

$$\Delta_{2'} = S \sqrt{\frac{T-1}{T}} \frac{\sigma_p^2}{\sqrt{\frac{\sigma_{2'}^2}{2} + \sigma_p^2}} .$$

For selection within blocks, the expected gain is the same as for the BIR design.

The normal equations necessary to solve for the $\hat{\beta}'$ s in the best index method of selection are as follows:

$$\frac{p-1}{pr} (\sigma^2 + r\sigma_p^2) \beta_{1'} + (0) \beta_{2'} = \frac{(p-1)}{p} \sigma_p^2$$

$$(0) \beta_{1'} + \frac{b-1}{bpr} (\sigma^2 + r\sigma_p^2 + p\sigma_{r'}^2 + pr\sigma_{b'}^2) \beta_{2'} = \frac{(b-1)}{bp} \sigma_p^2 .$$

From these equations, the β' s are:

$$\beta_{1'} = \frac{\sigma_p^2}{\sigma_p^2 + \frac{\sigma^2}{r}} = \beta_1$$

$$\beta_{2'} = \frac{\sigma_p^2}{(p\sigma_{b'}^2 + \frac{\sigma_{r'}^2}{r} + \sigma_p^2 + \frac{\sigma^2}{r})} .$$

The expected gain from selection for the best index is

$$\Delta_{I'} = S\sigma_p^2 \sqrt{\frac{r}{p} \left[\frac{p-1}{\sigma^2 + r\sigma_p^2} + \frac{b-1}{b(\sigma^2 + r\sigma_p^2 + p\sigma_r^2 + pr\sigma_b^2)} \right]} .$$

If check plots are included in the experiment and entry means are adjusted for check performance by the methods outlined in Section 7.2, the expected gains from selection are the same as for the BIR design, that is

$$\Delta_{1c'} = \Delta_{1c}$$

$$\Delta_{2c'} = \Delta_{2c} .$$

8.4 Comparisons of Methods of Selection

A measure of the relative efficiencies, in terms of expected gain, for the various methods of utilizing block information outlined in Section 7.2 may be obtained by expressing these Δ 's as fractions of $\Delta_{2'}$, the expected gain when block information is ignored. However, since $\Delta_{2'}$ is not independent of block size as Δ_2 was, the size of this ratio is not a direct reflection of the way in which these Δ 's vary with block size. With the exception of the best index, all of the expressions for expected gain for methods of selection involving block adjustments are the same for both the BIR and RIB designs. Since comparisons of these Δ 's as fractions of Δ_2 have already been made in Section 7.5, it is sufficient to look at the ratios of Δ_2 to $\Delta_{2'}$ and $\Delta_{I'}$ to $\Delta_{I'}$, to determine the effect of the RIB design on these comparisons.

The ratio of expected gain from selection ignoring blocks for the BIR design to the expected gain for the RIB design is

$$\frac{\Delta_2}{\Delta_{2'}} = \sqrt{\frac{\frac{T(b-1)}{b(T-1)} (r\sigma_{b'}^2 + \sigma_{r'}^2) + \sigma^2 + r\sigma_p^2}{\sigma_T^2 + r\sigma_p^2}} .$$

In terms of average error variances the ratio is

$$\frac{\Delta_2}{\Delta_{2'}} = \sqrt{\frac{\bar{\sigma}_{2'}^2 + 2\sigma_p^2}{\bar{\sigma}_2^2 + 2\sigma_p^2}} .$$

In this case, the ratio of expected gains is a monotonic function of the ratio of the average error variances and is always ≥ 1 since $\bar{\sigma}_{2'}^2 \geq \bar{\sigma}_2^2$.

An indication of the size of this ratio is achieved by expressing the environmental components of variance in terms of Smith's variance relationship. Values for the ratio $\Delta_2/\Delta_{2'}$, for $T = 256$ and $r = 2$ are given in Tables 14 and 15. While the expected gain from selection ignoring blocks is always as great or greater for the BIR design than for the RIB design, it is only for small B values and small block sizes that the difference becomes sizeable.

The ratios of Δ_I to $\Delta_{I'}$, are presented in Tables 16 and 17. The BIR design is again always equal to or better than its alternative, although the differences are very small over the entire ranges of B and b .

Table 14. Ratio of the expected gain from selection ignoring blocks for the BIR design to the expected gain for the RIB design, Δ_2/Δ_2^* , for $T=256$, $r=2$, and σ_p^2/σ_T^2 , approaching zero

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	1.061	1.042	1.017	1.006	1.002	1.000	1.000
4	1.117	1.084	1.037	1.014	1.004	1.001	1.000
8	1.173	1.127	1.061	1.025	1.008	1.001	1.000
16	1.225	1.172	1.090	1.040	1.014	1.003	1.000
32	1.275	1.219	1.126	1.062	1.025	1.005	1.000
64	1.323	1.266	1.167	1.092	1.041	1.010	1.000
128	1.369	1.316	1.216	1.133	1.068	1.019	1.000

Table 15. Ratio of the expected gain from selection ignoring blocks for the BIR design to the expected gain for the RIB design, Δ_2/Δ_2^* , for $T=256$, $r=2$, and $\sigma_p^2/\sigma_T^2 = 0.5$

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	1.029	1.020	1.008	1.003	1.001	1.000	1.000
4	1.057	1.041	1.018	1.007	1.002	1.000	1.000
8	1.085	1.063	1.030	1.012	1.004	1.001	1.000
16	1.112	1.086	1.045	1.020	1.007	1.001	1.000
32	1.138	1.110	1.063	1.031	1.012	1.003	1.000
64	1.163	1.135	1.085	1.047	1.021	1.005	1.000
128	1.188	1.162	1.111	1.068	1.034	1.009	1.000

Table 16. Ratio of the expected gain from selection on the best index for the BIR design to the expected gain for the RIB design, Δ_I/Δ_{I^*} , for $T=256$, $r=2$, and σ_p^2/σ_T^2 approaching zero

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000
8	1.000	1.000	1.001	1.001	1.001	1.001	1.000
16	1.001	1.001	1.002	1.002	1.002	1.002	1.000
32	1.003	1.003	1.005	1.006	1.006	1.003	1.000
64	1.009	1.010	1.014	1.016	1.015	1.007	1.000
128	1.032	1.037	1.047	1.048	1.039	1.016	1.000

Table 17. Ratio of the expected gain from selection on the best index for the BIR design to the expected gain for the RIB design, Δ_I/Δ_{I^*} , for $T=256$, $r=2$, and $\sigma_p^2/\sigma_T^2 = 0.5$

No. of blocks (b)	Regression coefficient (B)						
	0	0.1	0.3	0.5	0.7	0.9	1.0
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	1.000	1.000	1.000	1.000	1.000	1.000	1.000
8	1.001	1.001	1.001	1.001	1.001	1.000	1.000
16	1.002	1.003	1.003	1.003	1.002	1.001	1.000
32	1.008	1.008	1.008	1.008	1.006	1.002	1.000
64	1.022	1.022	1.021	1.018	1.012	1.004	1.000
128	1.067	1.063	1.054	1.041	1.026	1.009	1.000

8.5 Effect of Replication of Entries on Expected Gain from Selection

Since the expressions for expected gain from selection within blocks are the same for both the BIR and RIB designs, the effect of replication is the same as described in Section 7.6 for the BIR design. However, for selection ignoring blocks somewhat different results are obtained. The chief difference is the marked effect of block size on the efficiency achieved by replication.

Values for the ratio Δ_1/Δ_2 are presented in Table 18 for an experiment of 512 plots, where Δ_1 is the expected gain from selection ignoring blocks with 2 replications and Δ_2 is the expected gain using 4 replications.

These results indicate that the use of additional replications has little advantage, even when selection is intense and the entry component of variance very small, if the proper block size is used. This is in contrast to the Δ_1/Δ_2 comparisons for the BIR design.

The expected gain from selection using the best index varies with replication as it did for the BIR design, i.e., $\Delta_1/\Delta_2 < 1$ only with more intensive selection when σ_p^2 is very small. In addition, the ratio of Δ_1 to Δ_2 is practically constant for all block sizes.

Table 18. Values for the ratio Δ_2/Δ_1 , for an experiment of 512 plots, where Δ_1 is the expected gain from selection ignoring blocks with 2 replications and Δ_2 is the expected gain using 4 replications

Top 25 entries selected						
No. of entries per block	$\sigma_p^2/\sigma_{TB}^2 \rightarrow 0$			$\sigma_p^2/\sigma_{TB}^2 = 0.2$		
	.3	.5	.7	.3	.5	.7
64	.882	.883	.886	1.000	1.003	1.006
16	.986	.959	.897	1.063	1.048	1.011
4	1.025	.967	.923	1.083	1.048	1.024
Top 51 entries selected						
64	1.019	1.020	1.023	1.156	1.159	1.162
16	1.139	1.108	1.036	1.228	1.211	1.168
4	1.184	1.117	1.067	1.252	1.211	1.183

9.0 THE SIMPLE LATTICE DESIGN

9.1 Model and Equations for Adjusted Entry Means

Although a large number of balanced and partially balanced incomplete block designs have been developed, many of them are unsuitable for genetic experiments involving a large number of entries, since the requirement of balance drastically restricts the range of block sizes and minimum number of replications that can be used. In response to the need for efficient designs for a large number of entries, Yates (1936) introduced a group of incomplete block designs known as quasi-factorials or lattices. Perhaps the simplest of these designs is the simple or double lattice which requires only two replications for partial balance and a total of p^2 entries.

In constructing lattice designs, the entries are designated in the manner used for factorial arrangements. For example, $p^2 = (2)^2$ entries are numbered 00, 01, 10, and 11, and even though the four entries are not a factorial arrangement of the factors a and b, they are comparable to a factorial arrangement for purposes of design and analysis.

If, in a simple lattice, the A pseudo effect is confounded with block differences in the first replicate and the B pseudo effect is confounded in the second replicate, the resulting entry arrangements before randomization are

Replicate I (Arrangement X)				Replicate II (Arrangement Y)			
Level of effect	Block		Mean	Level of effect	Block		Mean
	1	2			1	2	
(A) ₀	t ₁	t ₂	x _{.11}	(B) ₀	t ₁	t ₃	y _{1.2}
	(00)	(01)			(00)	(10)	
(A) ₁	t ₃	t ₄	x _{.21}	(B) ₁	t ₂	t ₄	y _{2.2}
	(10)	(11)			(01)	(11)	
Mean	x _{1.1}	x _{2.1}	x _{..1}	Mean	y _{.12}	y _{.22}	y _{..2}

The field design is the same as the BIR design except entries are regrouped into the blocks in each replication.

To simplify the notation, Federer (1955) suggested the use of two equations to represent the model for such an experiment:

$$X_{ij1} = u + r_1 + b_{i1} + p_{ij} + e_{ij1} \quad (9.1.1)$$

$$Y_{ij2} = u + r_2 + b_{j2} + p_{ij} + e_{ij2} \quad , \quad (9.1.2)$$

where u = mean effect

b_{i1} = incomplete block effect in the X arrangement

b_{j2} = incomplete block effect in the Y arrangement

r_k = replicate effect in the kth replicate

p_{ij} = entry effect

e_{ijk} = random effect .

The subscripts i and j vary from one to p . All effects are assumed to be random and uncorrelated and in addition it is assumed that the block and replication effects are the same as defined for Model 7.1.3 of the

BIR design. Thus, σ^2 and σ_b^2 may be written in terms of Smith's parameter B as in equations 7.1.5 and 7.1.6.

When lattice designs were first introduced, attention was directed to methods for extracting the intra-block information which gives estimates of entry differences based on comparisons arising within a block. Although it was recognized that comparisons of block means or block totals contained some information on entry comparisons, they were neglected on the grounds that such comparisons are subject to greater environmental variability.

If inter-block information is ignored, a least squares estimate of the entry effect may be obtained as for complete block designs by minimizing

$$\sum_i \sum_j [(X_{ij1} - u - r_1 - b_{i1} - p_{ij})^2 + (Y_{ij2} - u - r_2 - b_{j2} - p_{ij})^2]$$

subject to the usual restrictions

$$\sum_i r_k = \sum_i b_{i1} = \sum_j b_{j2} = \sum_i \sum_j p_{ij} = 0.$$

Under these conditions,

$$\hat{p}_{ij} = \frac{1}{2}(X_{ij1} + Y_{ij2} - x_{..1} - y_{..2}) + \frac{1}{2}(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2}) \quad (9.1.3)$$

and the least squares estimate of the adjusted entry mean is

$$\hat{Y}_{ij.} = \hat{p}_{ij} + \hat{u} = \frac{1}{2}(X_{ij1} + Y_{ij2}) + \frac{1}{2}(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2}). \quad (9.1.4)$$

That is, the unadjusted mean, $(\frac{1}{2})(X_{ij1} + Y_{ij2})$, is adjusted for the correction for block means, $(\frac{1}{2})(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2})$.

When inter-block information is recovered, it is helpful to utilize the pseudo-factorial arrangement of the entries in deriving the adjusted value of an entry mean. For the simple example of $p^2 = 4$ entries, the estimate of the A effect is

$$\frac{1}{2}(01 + 11 - 00 - 10) .$$

The estimate from replicate II in which the A effect is unconfounded is a within block estimate with error variance σ^2 , and the estimate from replicate I where the effect is confounded has an error variance of $(\sigma^2 + 2\sigma_b^2)$. The combined estimate with the least variance is

$$\hat{A} = \frac{W'A_{II} + WA_{I}}{W' + W} = \frac{1}{2}(A_{II} + A_I) + \frac{W' - W}{2(W' + W)} (A_{II} - A_I) ,$$

where $W = \frac{1}{\sigma^2}$

$$W' = \frac{1}{\sigma^2 + 2\sigma_b^2}$$

A_{II} = estimate of the A effect from replicate II

A_I = estimate of the A effect from replicate I .

Similarly,

$$\hat{B} = \frac{W'B_{II} + WB_I}{W' + W} = \frac{1}{2}(B_{II} + B_I) + \frac{W' - W}{2(W' + W)} (B_{II} - B_I)$$

and

$$\hat{AB} = \frac{1}{2}(AB_I + AB_{II}) .$$

Note that the AB interaction effect is not confounded in either replicate.

The adjusted value of any one entry, say t_1 , is given by

$$\hat{t}_1 = 00 - \hat{u} - \frac{\hat{A}}{2} - \frac{\hat{B}}{2} + \frac{\hat{AB}}{2} .$$

Substituting in values for $\hat{u}$, $\hat{A}$, $\hat{B}$, and $\hat{AB}$ and generalizing the equation to p levels of the factors a and b results in

$$\hat{Y}_{ij.} = \frac{1}{2}(X_{ij1} + Y_{ij2}) + \frac{w-w'}{2(w+w')} (y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2}) \quad (9.1.5)$$

where $w = W = \frac{1}{\sigma^2}$

$$w' = \frac{1}{\sigma^2 + p\sigma_b^2}.$$

Thus, when inter-block information is recovered, the size of the block correction is dampened by the weighting factor, $\frac{w-w'}{w+w'}$.

These results can be put in the form of the general prediction equation of Section 7.2 by writing

$$\begin{aligned} \hat{Y}_{ij.} = \frac{1}{2}[x_{..1} + y_{..2} + \hat{\beta}_{1L}(X_{ij1} + Y_{ij2} - x_{..1} - y_{..2}) \\ + \hat{\beta}_{2L}(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2})] \end{aligned} \quad (9.1.6)$$

and since only the relative magnitude of the $\hat{\beta}$'s is of importance in ranking the entries, 9.1.6 may be expressed simply as

$$\hat{Y}_{ij.} = \frac{1}{2}[X_{ij1} + Y_{ij2} + \beta(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2})] \quad (9.1.7)$$

where

$$\beta = \frac{\hat{\beta}_{2L}}{\hat{\beta}_{1L}}.$$

If blocking is ignored, the $\hat{\beta}$'s take the values

$$\hat{\beta}_{1L} = 1, \hat{\beta}_{2L} = 0 \quad (\beta = 0)$$

and the predicted value of an entry mean is

$$\hat{y}_{ij.} = \frac{1}{2}(x_{ij1} + y_{ij2}).$$

For within blocks selection (inter-block information is ignored), the $\hat{\beta}$'s are

$$\hat{\beta}_{1L} = \hat{\beta}_{2L} = 1 \quad (\beta = 1) .$$

For the index method of selection, $\hat{\beta}$'s are chosen for equation 9.1.6 that will maximize the correlation between the true values and predicted values of entry means. These results will be given in Section 9.3.

If block means only are used in the intra-block adjustment and no use is made of the pseudo-factorial arrangement of the treatments, the adjusted entry mean is

$$\hat{y}_{ij.} = \frac{1}{2}(x_{..1} + y_{..2}) + \frac{1}{2}(x_{ij1} + y_{ij2} - x_{i.1} - y_{.j2}). \quad (9.1.8)$$

9.2 Average Error Variance of Entry Comparisons

Rao (1947) described a general method of analysis for incomplete block designs in which a system of associates was set up based on the "association" of entries in the incomplete block. First associates are entries which do not appear together in an incomplete block and second associates refer to pairs of entries appearing together once in an incomplete block.

In the simple lattice design, an entry appears with $(p-1)$ other entries in an incomplete block in both the X and Y arrangements. Thus,

each of the p^2 entries occurs with $2(p-1)$ entries in an incomplete block in the two groupings. Likewise, there are $(p-1)^2$ entries with which a given entry is not associated in an incomplete block. However, it must be remembered that in each case, half of the comparisons are duplicates. Thus, there is a total of $\frac{p^2(p-1)^2}{2}$ comparisons among first associates and $\frac{2(p-1)p^2}{2}$ comparisons among second associates.

When inter-block information is not included in the analysis of a simple lattice experiment, Yates (1936) found that the variance among first associate comparisons is $\sigma^2 \frac{(p+2)}{p}$ and among second associate comparisons is $\sigma^2 \frac{(p+1)}{p}$. Therefore, the average error variance per comparison is

$$\begin{aligned}\sigma_{1L}^2 &= \frac{\frac{p^2(p-1)^2}{2} \frac{(p+2)}{p} \sigma^2 + \frac{2(p-1)p^2}{2} \frac{(p+1)}{p} \sigma^2}{\frac{p^2(p-1)(p+1)}{2}} \\ &= \frac{(p+3)}{p+1} \sigma^2.\end{aligned}$$

If inter-block information is included as in equation 9.1.5, the variance among first associate comparisons is

$$\frac{2}{p} \left(\frac{2}{w+w'} + \frac{p-2}{2w} \right)$$

and among second associate comparisons is

$$\frac{2}{p} \left(\frac{1}{w+w'} + \frac{p-1}{2w} \right).$$

The average error variance per comparison is

$$\sigma_{\beta*L}^2 = \frac{2}{p+1} \left(\frac{2}{w+w'} + \frac{p-1}{2w} \right) = \frac{2\sigma^2}{p+1} \left[\frac{2(\sigma^2 + p\sigma_b^2)}{2\sigma^2 + p\sigma_b^2} + \frac{p-1}{2} \right] \dots$$

Note that the quantity in brackets can never become greater than $\frac{(p+3)}{2}$. Therefore, in general, σ_{1L}^2 is greater than $\sigma_{\beta*L}^2$, but $\sigma_{\beta*L}^2$ requires that w and w' be known.

When entry means are adjusted only for block means (equation 9.1.8), the variance among first associate comparisons is

$$\frac{(p-1)}{p} \sigma^2 + \frac{(p-2)}{p^2} \sigma_p^2$$

and among second associate comparisons is

$$\frac{(2p-1)}{2p} \sigma^2 + \frac{(p-1)}{2p^2} \sigma_p^2.$$

The average error variance per comparison is

$$\sigma_{2L}^2 = \left[\frac{p(p-1)^2 + (2p-1)}{p^2(p+1)} \right] \sigma^2 + \frac{(p-1)^2}{p^2(p+1)} \sigma_p^2.$$

Thus, the size of σ_{2L}^2 relative to σ_{1L}^2 and $\sigma_{\beta*L}^2$ is influenced by the magnitude of σ_p^2 .

9.3 Expected Gain from Selection

In order to formulate the expected gain from selection on the adjusted entry means, it is necessary to derive the variances and covariances of the deviations in equations 9.1.6, 9.1.7, and 9.1.8 and the covariances between the deviations and the true values of entry means.

The following expectations are utilized in obtaining these variances and covariances:

$$\begin{aligned} E(X_{ij1})(y_{i.2}) &= E(X_{ij1})(y_{.j2}) = E(Y_{ij2})(x_{i.1}) = E(Y_{ij2})(x_{.j1}) \\ &= E(x_{i.1})(y_{i.2}) = E(x_{.j1})(y_{.j2}) = u^2 + \frac{\sigma_p^2}{p} \end{aligned}$$

$$\begin{aligned} E(x_{..1})(y_{..2}) &= E(X_{ij1})(y_{..2}) = E(Y_{ij2})(x_{..1}) = E(y_{i.2})(x_{..1}) \\ &= E(x_{i.1})(y_{..2}) = E(x_{.j1})(y_{..2}) = E(y_{.j2})(x_{..1}) \\ &= E(x_{.j1})(y_{i.2}) = E(x_{i.1})(y_{.j2}) = u^2 + \frac{\sigma_p^2}{p^2} \end{aligned}$$

$$\begin{aligned} E(X_{ij1})(x_{i.1}) &= E(Y_{ij2})(y_{.j2}) = E(x_{i.1})^2 = E(y_{.j2})^2 \\ &= u^2 + \sigma_r^2 + \sigma_b^2 + \frac{\sigma_p^2}{p} + \frac{\sigma_p^2}{p} \end{aligned}$$

$$\begin{aligned} E(X_{ij1})(x_{.j1}) &= E(Y_{ij2})(y_{i.2}) = E(y_{i.2})^2 = E(x_{.j1})^2 \\ &= u^2 + \sigma_r^2 + \frac{\sigma_b^2}{p} + \frac{\sigma_p^2}{p} + \frac{\sigma_p^2}{p} \end{aligned}$$

$$\begin{aligned} E(X_{ij1})(x_{..1}) &= E(y_{.j2})(y_{i.2}) = E(x_{.j1})(x_{i.1}) = E(Y_{ij2})(y_{..2}) \\ &= E(y_{i.2})(y_{..2}) = E(x_{i.1})(x_{..1}) = E(x_{.j1})(x_{..1}) \\ &= E(y_{.j2})(y_{..2}) = E(x_{..1})^2 = E(y_{..2})^2 \\ &= u^2 + \sigma_r^2 + \frac{\sigma_b^2}{p} + \frac{\sigma_p^2}{p^2} + \frac{\sigma_p^2}{p^2} \end{aligned}$$

$$E(X_{ij1})^2 = E(Y_{ij2})^2 = u^2 + \sigma_r^2 + \sigma_b^2 + \sigma_p^2 + \sigma^2$$

$$E(X_{ij1})(Y_{ij2}) = u^2 + \sigma_p^2$$

$$E p(y_{i.2}) = E p(x_{i.1}) = E p(x_{.j1}) = E p(y_{.j2}) = \frac{\sigma_p^2}{p}$$

$$E p(X_{ij1}) = E p(Y_{ij2}) = \sigma_p^2$$

$$E p(x_{..1}) = E p(y_{..2}) = \frac{\sigma_p^2}{p}$$

The variances and covariances of the deviations are

$$V \left\{ \frac{1}{2} [(X_{ij1} + Y_{ij2} - x_{..1} - y_{..2}) + \beta(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2})] \right\}$$

$$= \frac{1}{4} E (X_{ij1} + Y_{ij2} - x_{..1} - y_{..2} + \beta y_{i.2} - \beta x_{i.1} + \beta x_{.j1} - \beta y_{.j2})^2$$

$$= \frac{p-1}{p^2} \left[\frac{p(\beta-1)^2}{2(p+1)} \sigma_b^2 + \frac{(2\beta^2+p+1)}{2(p+1)} \sigma^2 + \sigma_p^2 \right]$$

$$V \left[\frac{1}{2} (y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2}) \right] = \frac{1}{4} E (y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2})^2$$

$$= \frac{p-1}{2p^2} (2\sigma^2 + p\sigma_b^2)$$

$$V \left[\frac{1}{2} (X_{ij1} + Y_{ij2} - x_{i.1} - y_{.j2}) \right] = \frac{1}{4} E (X_{ij1} + Y_{ij2} - x_{i.1} - y_{.j2})^2$$

$$= \frac{p-1}{2p} \left[\frac{(2p-1)}{p} \sigma_p^2 + \sigma^2 \right]$$

$$V \left[\frac{1}{2} (X_{ij1} + Y_{ij2} - x_{..1} - y_{..2}) \right] = \frac{1}{4} E (X_{ij1} + Y_{ij2} - x_{..1} - y_{..2})^2$$

$$= \frac{p-1}{p} \left[\frac{\sigma_b^2}{2} + \frac{(p+1)}{p} \sigma_p^2 + \frac{(p+1)}{2p} \sigma^2 \right]$$

$$\text{Cov} \left[\frac{1}{2} (X_{ij1} + Y_{ij2} - x_{..1} - y_{..2}) \right] \left[\frac{1}{2} (y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2}) \right]$$

$$= -\frac{(p-1)}{2p} \sigma_b^2$$

$$\begin{aligned}
& \text{Cov } p \left\{ \frac{1}{2} [(X_{ij1} + Y_{ij2} - x_{..1} - y_{..2}) + \beta(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2})] \right\} \\
&= \text{Cov } p \left[\frac{1}{2} (X_{ij1} + Y_{ij2} - x_{..1} - y_{..2} + y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2}) \right] \\
&= \frac{(p^2 - 1)}{p^2} \sigma_p^2
\end{aligned}$$

$$\text{Cov } p \left[\frac{1}{2} (X_{ij1} + Y_{ij2} - x_{i.1} - y_{.j2}) \right] = \frac{(p-1)}{p} \sigma_p^2$$

$$\text{Cov } p \left[\frac{1}{2} (y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2}) \right] = 0.$$

The expected gain from selecting among entry means adjusted for both intra-block and inter-block information is

$$\begin{aligned}
\Delta_{\beta L} &= \frac{S \text{Cov } p \left\{ \frac{1}{2} [(X_{ij1} + Y_{ij2} - x_{..1} - y_{..2}) + \beta(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2})] \right\}}{\sqrt{V \left\{ \frac{1}{2} [(X_{ij1} + Y_{ij2} - x_{..1} - y_{..2}) + \beta(y_{i.2} - x_{i.1} + x_{.j1} - y_{.j2})] \right\}}} \\
&= S \sqrt{\frac{p^2 - 1}{p^2}} \frac{\sigma_p^2}{\sqrt{\frac{p(\beta-1)^2}{2(p+1)} \sigma_b^2 + \frac{(2\beta^2 + p + 1)}{2(p+1)} \sigma^2 + \sigma_p^2}}
\end{aligned}$$

and when $\beta = \beta^*$, $\Delta_{\beta L}$ reduces to

$$\Delta_{\beta^* L} = S \sqrt{\frac{p^2 - 1}{p^2}} \frac{\sigma_p^2}{\sqrt{\left(\frac{\beta^*}{p+1} + \frac{1}{2}\right) \sigma^2 + \sigma_p^2}}.$$

For the special case where $\beta=1$ (inter-block information is ignored), the expected gain is

$$\Delta_{1L} = S \sqrt{\frac{p^2 - 1}{p^2}} \frac{\sigma_p^2}{\sqrt{\frac{(p+3)}{2(p+1)} \sigma^2 + \sigma_p^2}}.$$

In terms of average error variances these expressions are

$$\Delta_{1L} = S \sqrt{\frac{p^2-1}{p^2}} \frac{\sigma_p^2}{\sqrt{\frac{\sigma_{1L}^2}{2} + \sigma_p^2}}$$

and

$$\Delta_{\beta^*L} = S \sqrt{\frac{p^2-1}{p^2}} \frac{\sigma_p^2}{\sqrt{\frac{\sigma_{\beta^*L}^2}{2} + \sigma_p^2}} .$$

If block means only are used in the intra-block adjustment and no use is made of the pseudo-factorial arrangement of the treatments, the expected gain from selection is

$$\begin{aligned} \Delta_{2L} &= \frac{S \text{Cov } p[\frac{1}{2}(X_{1j1} + Y_{1j2} - x_{1.1} - y_{.j2})]}{\sqrt{V[\frac{1}{2}(X_{1j1} + Y_{1j2} - x_{1.1} - y_{.j2})]}} \\ &= S \sqrt{\frac{p-1}{p}} \frac{\sigma_p^2}{\sqrt{\frac{\sigma^2}{2} + (1 - \frac{1}{2p}) \sigma_p^2}} . \end{aligned}$$

The normal equations necessary to solve for the index estimates of the $\hat{\beta}$'s in equation 9.1.6 are

$$\begin{aligned} \frac{p-1}{p} \left[\frac{\sigma_b^2}{2} + \frac{(p+1)}{p} \sigma_p^2 + \frac{(p+1)}{2p} \sigma^2 \right] \beta_{1L} - \left[\frac{(p-1)}{p} \frac{\sigma_b^2}{2} \right] \beta_{2L} &= \frac{(p^2-1)}{p^2} \sigma_p^2 \\ - \left[\frac{(p-1)}{p} \frac{\sigma_b^2}{2} \right] \beta_{1L} + \frac{p-1}{p} \left[\frac{\sigma_b^2}{2} + \frac{\sigma^2}{p} \right] \beta_{2L} &= 0 . \end{aligned}$$

From the second equation, the relationship between the β 's is

$$\beta_{1L} = \frac{(2\sigma^2 + p\sigma_b^2)}{p\sigma_b^2} \beta_{2L}$$

and

$$\beta = \frac{\beta_{2L}}{\beta_{1L}} = \frac{p\sigma_b^2}{2\sigma^2 + p\sigma_b^2} = \beta^* .$$

Thus, the weight assigned to the block correction in the index is equivalent to the estimate obtained in Section 9.1 when both intra-block and inter-block information was utilized in estimating factorial effects, and the expected gain for the index is Δ_{β^*L} .

9.4 Comparisons of Methods of Selection

An indication of the merits of the various methods of adjusting entry means may be obtained by examining the ratios of expected gains. The ratio of the expected gain from selection among entry means adjusted for both inter-block and intra-block information to the expected gain for the intra-block adjustment only is

$$\frac{\Delta_{\beta L}}{\Delta_{1L}} = \sqrt{\frac{\frac{(p+3)}{p+1} \sigma^2 + 2\sigma_p^2}{\frac{p(\beta-1)^2}{p+1} \sigma_b^2 + \frac{(2\beta^2 + p + 1)}{p+1} \sigma^2 + 2\sigma_p^2}} .$$

This ratio is a maximum when $\beta = \beta^*$, the least squares value obtained previously. The β values for which $\Delta_{\beta L} = \Delta_{1L}$ may be obtained by solving the quadratic equation

$$\beta^2(pC+2) - 2\beta pC + pC - 2 = 0$$

where $C = \frac{\sigma_b^2}{\sigma^2}$.

Thus,

$$\beta = \frac{pC + 2}{pC + 2} = 1, \frac{pC-2}{pC+2}$$

and $\Delta_{\beta L} > \Delta_{1L}$ when

$$\frac{pC-2}{pC+2} < \beta < 1 .$$

When $\beta = \beta^*$ the ratio of expected gains is

$$\frac{\Delta_{\beta^* L}}{\Delta_{1L}} = \sqrt{\frac{\frac{(p+3)}{p+1} \sigma^2 + 2\sigma_p^2}{(1 + \frac{2\beta^*}{p+1}) \sigma^2 + 2\sigma_p^2}} = \sqrt{\frac{\sigma_{1L}^2 + 2\sigma_p^2}{\sigma_{\beta^* L}^2 + 2\sigma_p^2}} .$$

An application of Smith's variance relationship permits a numerical evaluation of the ratio. For an experiment of $p^2 = 256$ entries, $B = 0.5$, and σ_p^2/σ_T^2 , approaching zero

$$\frac{\Delta_{\beta^* L}}{\Delta_{1L}} = 1.022$$

and when $\sigma_p^2/\sigma_T^2 = 0.5$

$$\frac{\Delta_{\beta^* L}}{\Delta_{1L}} = 1.010 .$$

The effect of the recovery of inter-block information on the average error variance of pair-wise comparisons is in agreement with this result, since $\sigma_{1L}^2 > \sigma_{\beta^* L}^2$.

The ratio of $\Delta_{\beta L}$ to Δ_{2L} , the expected gain when pseudo-factorial information is not utilized, is

$$\frac{\Delta_{BL}}{\Delta_{2L}} = \sqrt{\frac{p+1}{p}} \sqrt{\frac{\sigma^2 + \frac{(2p-1)}{p} \sigma_p^2}{\frac{p(\beta-1)^2}{p+1} \sigma_b^2 + \frac{(2\beta^2 + p+1)}{p+1} \sigma^2 + 2\sigma_p^2}} .$$

The β value that maximizes the ratio is the least squares estimate as before, but the β values for which $\Delta_{BL} > \Delta_{2L}$ are much more complicated than the previous expressions and are functions of σ_p^2 as well as σ^2 and σ_b^2 . For $p^2 = 256$ entries, $\beta = \beta^*$, $B = 0.5$, and σ_p^2/σ_T^2 , approaching zero,

$$\frac{\Delta_{\beta^*L}}{\Delta_{2L}} = 0.996$$

and when $\sigma_p^2/\sigma_T^2 = 0.5$

$$\frac{\Delta_{\beta^*L}}{\Delta_{2L}} = 1.006 .$$

Note that $\Delta_{\beta^*L} < \Delta_{2L}$ when σ_p^2 is small. Therefore, the index utilized is not the best index possible, since the expected gain for the best index is theoretically always greater than for any other method of selection. The ratio of average error variances is also affected by the size of σ_p^2 and when $\sigma_p^2/\sigma_T^2 = 0$ is

$$\frac{\sigma_{2L}^{-2}}{\sigma_{\beta^*L}^{-2}} = 0.779 .$$

When $\sigma_p^2/\sigma_T^2 = 0.5$,

$$\frac{\sigma_{2L}^{-2}}{\sigma_{\beta^*L}^{-2}} = 0.808 .$$

The ratio of Δ_{1L} to Δ_{2L} is

$$\frac{\Delta_{1L}}{\Delta_{2L}} = \sqrt{\frac{p+1}{p}} \sqrt{\frac{\sigma^2 + \frac{(2p-1)}{p} \sigma_p^2}{\frac{(p+3)}{p+1} \sigma^2 + 2\sigma_p^2}}$$

and $\Delta_{1L} > \Delta_{2L}$ when $\frac{\sigma_p^2}{\sigma^2} > \frac{p}{p+1}$.

It is also of interest to compare Δ_{1L} and Δ_{2L} with Δ_1 , the expected gain from within blocks selection for the BIR and RIB designs. The ratio of Δ_{1L} to Δ_1 is

$$\frac{\Delta_{1L}}{\Delta_1} = \sqrt{\frac{p+1}{p}} \sqrt{\frac{\sigma^2 + 2\sigma_p^2}{\frac{(p+3)}{p+1} \sigma^2 + 2\sigma_p^2}}.$$

In this case $\Delta_{1L} > \Delta_1$ when $\frac{\sigma_p^2}{\sigma^2} > \frac{p-1}{2(p+1)}$.

The ratio of Δ_{2L} to Δ_1 is

$$\frac{\Delta_{2L}}{\Delta_1} = \sqrt{\frac{\sigma^2 + 2\sigma_p^2}{\sigma^2 + \frac{(2p-1)}{p} \sigma_p^2}}$$

and $\Delta_{2L} > \Delta_1$ except when $\sigma_p^2 = 0$.

In general, the expected gain from selection employing block information only, Δ_{2L} , is as good or better than the expected gain from the index employing inter-block information, Δ_{B*L} , as well as the least squares method of intra-block adjustment, Δ_{1L} , except when σ_p^2 is large. It is also as good or better (when $\sigma_p^2 \neq 0$) than selecting within blocks in the BIR and RIB designs. When blocking is ignored, the expected gain for the simple lattice design is the same as for the BIR design.

10.0 GENERAL DISCUSSION AND CONCLUSIONS

Although the incomplete block designs considered in this study admittedly do not exhaust all the possibilities, they are very useful for the evaluation of genetic material in that the number of replications and block sizes required are relatively unrestricted. The BIR and RIB designs are most versatile in this respect since no restrictions are imposed. The simple lattice design requires two replications for partial balance and a block size equal to the square root of the total number of entries. If a completely balanced lattice design is desired, a total of $(p+1)$ replications is required. For other balanced incomplete block designs, the number of replications required varies with block size and the total number of entries. Although balanced designs have the desirable property of measuring all entry comparisons with equal precision, the number of replications required often makes the use of such designs prohibitive in genetic selection experiments where the number of entries is usually large.

It must be recognized at the outset that the usefulness of the numerical values of expected gain for the examples outlined in the preceding sections rests on the assumption that Smith's variance law, 6.0.1, is appropriate. Since the law implies that adjacent areas in a field are equally correlated, the relationship has limitations.

A further limitation is encountered when these results are applied to actual field experiments in that the true value of the regression coefficient, B , is not known and must be estimated. The majority of the estimates obtained by Smith (1938), Brim and Mason (1959),

Robinson et al (1948) and others fall in the range from 0.4 to 0.7. Fortunately, the relative sizes of the estimates of expected gain do not change very rapidly for B values in this range so that the various methods of selection can be compared in a general way in this range without knowing the exact value of B.

Although the best index method of selection is theoretically always equal to or better than any of the other alternative methods, in practice, the true index weights are not known. The bounds on the optimum weights for the two-variate indexes considered are $0 \leq \beta_2 \leq \beta_1 \leq 1$. Patel (1962) found that if the estimated values of the weights are restricted to the same bounds, the estimated index is in general always better than the base index consisting of unadjusted entry means. However, when optimum block size is used the comparisons in this study indicate that selecting within blocks is also better than the base index and is only slightly inferior to the optimum index. The index employed in the simple lattice design is not the optimum index. It is equivalent to the inter-block least squares method and is always equal to or better than the least squares intra-block method of adjusting entry means (9.1.4). It is less efficient, when σ_p^2 is small, than adjusting entry means for block effects only (9.1.8). The latter method of selection is always equal to or better than selecting within blocks in the RIB and BIR designs.

When blocking is ignored, the expected gain from selection is the same for the BIR and simple lattice designs and is always equal to or greater than the expected gain for the RIB design. The expected gain for selecting within blocks is the same for both the BIR and RIB designs.

In general, the inclusion of check plots in an incomplete block design for the purpose of adjusting entry means does not appear to be a desirable alternative. Yates (1936) came to a similar conclusion on the basis of comparisons of the average error variances involved. However, it must be recognized that the inclusion of checks for other comparative purposes may be desirable. An example is the situation in which selections are tested over time and a check or standard is included as a means of evaluating progress in the selected population. In such cases, the most effective means of utilizing the check information, except for the selection index approach, is by adjusting entry means for block means of both entries and checks with the block means weighted in proportion to the number of entries and checks included in the block (7.2.5).

Adjusting entry means for block deviations appears to be the most desirable way of ranking the entries, in a practical sense, in all of the designs considered, providing that the block size is near optimum. For a given block size, the simple lattice design is slightly more efficient than the BIR and RIB designs in this respect. However, this does not imply that the simple lattice is always the best design since the block size used in a lattice design is restricted to the square root of the total number of entries. If this particular block size is not the one which maximizes the expected gain from selection within blocks, the BIR and RIB designs might be the best choice. The examples of 256 entries cited in previous sections indicate optimum block size for such an experiment consists of about 16 plots if B is assumed to be in the

neighborhood of one half. In this particular case, the optimum block size for selection within blocks is compatible with the block size required for a simple lattice design.

From the standpoint of expected gain, there seems to be little reason for using the RIB design in a selection experiment since the BIR design is always equal to or better than the RIB design for the various methods of selection. However, the differences in expected gains are not large so that the RIB design might be resorted to in instances where its use would facilitate experimental operations such as planting and harvesting.

The comparisons made in Sections 7.6 and 8.5 indicate that, in general, the replication of entries in a selection experiment is of no advantage when the addition of replications results in fewer entries tested. However, most plant breeders prefer to use at least two replications to permit estimation of the error variance.

The inclusion of family structure in an incomplete block design presents some additional complications. If half sib and full sib families are included in the design, as illustrated in Section 7.7, the genetic covariance becomes a function of half sib and full sib genetic correlations and the entry component of variance is subdivided into contributions from both the maternal and paternal parents. Thus, the problem of determining optimum block size is made more complex by the modifying effects of family structure.

For selection on full sib family means within blocks and ignoring blocks, maximum expected gain is realized when unrelated entries are

included in the experiment. The expected gain from selecting within blocks declines rapidly as more and more half sib families are included in a block. For selection ignoring blocks, the effect is much less pronounced since entries in different blocks are always assumed to be unrelated, and for a given family structure the expected gain is independent of block size. If selection is based on half sib family means, the effect of family structure is somewhat different and is much dependent on size of block, there being an optimum family structure for each size of block within which selection is practiced. If blocking is ignored, the expected gain for a particular family structure is not independent of block size as it was for full sib family selection.

The information available from full sib, half sib, and block means may be combined in various ways to form selection indices. If half sib information is not included in an index, the expected gain is always maximum for unrelated full sibs, but when half sib information is given its appropriate weight, the maximum is obtained from an optimum family structure. For $b=1$ in the example of 64 entries considered in Section 7.7, this optimum occurs when the 64 full sib families fall into 8 half sib families. For $b > 1$ the optimum varies from 1 to 4 half sib families per block depending on block size. An index consisting of full sib and block information is only slightly better than full sib selection ignoring blocks, but, for small block sizes, is much better than selecting within blocks.

The complete index consisting of full sib, half sib, and block information is better than any of the reduced indices, particularly for

small block sizes. However, for small block sizes, the relative advantage of the complete index over full sib selection ignoring blocks is not very great.

It should be pointed out that the effect of blocking and family structure on the expected gain from selection is not the only consideration, since the simultaneous estimation of genetic variances is often desired. From this standpoint, family structure is a necessity. Blocking may be desirable since it permits distribution of the degrees of freedom more evenly among the mean squares as well as reduction in the error variance. The previous comparisons would indicate that for experiments grown for the joint purpose of estimation and selection, the best method of selection to use is highly dependent upon block size and the degree of family structure employed. Thus, no over-all recommendation can be made, although, in general, any within block method of selection becomes less efficient as the ratio of half sibs to full sibs within a block increases beyond an optimum and selection ignoring blocks becomes relatively more efficient, particularly for small block sizes.

It is of significance to note that in all situations where pairwise entry comparisons involve entry effects other than the two being compared, the relative efficiency based on average error variance is not monotonically related to the relative efficiency computed on expected gains. This occurs when entry means are adjusted for block effects in the BIR and RIB designs. Thus, the block size that optimizes relative efficiency in terms of average error variances does not necessarily optimize the relative efficiency based on expected gains. When entry

means are adjusted for the block correction in the simple lattice design by either the inter- or intra-block least squares methods, the average error variances are monotonically related to the expected gains. However, if the pseudo-factorial arrangement of the entries is ignored in adjusting for block effects this is no longer true, and the average error variance contains an entry component.

In some respects, the developments presented are over simplified. For example, the effect of testing at different locations and in different years has not been considered in formulating expected gain and no allowance has been made for genotype by environmental interactions. It must be recognized that such effects would be an over-riding factor, particularly if they accounted for a large part of the total variance, and would generally decrease the relative magnitude of the differences in expected gains for the designs compared. However, these effects should not switch the superiority of any of the designs since they affect each of the expected gains in a similar fashion. Other possibilities arise for the lattice design when more environments are included so that additional replicates can be blocked to give more balance in the total set of replications. This leads to further relative superiority of the lattice designs except that the genotype by environment variance will be somewhat of an over-riding factor. For a balanced lattice achieved in this way the average error variance of the least squares intra-block comparisons is $\frac{(p+1)^2}{p(p+3)}$ times that of a simple lattice replicated a like number of times. The increase in efficiency for the inter-block least squares method of selection is of a similar nature.

11.0 SUMMARY

An important choice that must be made by a plant breeder is an environmental design for testing his genetic entries in order to select among them. The purpose of this study was to compare alternative designs or procedures as to their gains from selection. Factors varied in the designs included number of check plots, number of replications, size of groups and family structure of entries. Considered were the block in replication (BIR), the replication in blocks (RIB), and the simple or double lattice designs.

To facilitate these comparisons and to clarify the components of environmental variance involved, an empirical variance relationship derived by Smith (1938) was used to relate block size to the variability between and within groups. Use of this relationship permitted a numerical comparison of the expected gains for various methods of selection and alternative blocking patterns.

For a given block size, the simple lattice design was more efficient than either the BIR or RIB designs and the BIR design was always equal to or better than the RIB design.

The method of selecting within blocks was almost as good as the best index and was considerably better than selection ignoring blocks, providing the optimum block size was used. The use of check plots and additional replications did not appear to be desirable alternatives, when they resulted in fewer entries tested.

The effects of family structure and genotype by environmental interactions on these results were discussed and some of the limitations of the methods used were pointed out.

12.0 LIST OF REFERENCES

- Brim, C. A. and Mason, D. D. 1959. Estimates of optimum plot size for soybean yield trials. *Agron. J.* 51:331-334.
- Cochran, W. G. and Cox, G. M. 1957. *Experimental Designs*. John Wiley and Sons, Inc. New York.
- Comstock, R. E. and Robinson, H. F. 1948. The components of genetic variance in populations. *Biometrics* 4:254-266.
- Federer, W. T. 1955. *Experimental Design*. The Macmillan Co. New York.
- Fisher, R. A. and Yates, F. 1953. *Statistical Tables*. Oliver and Boyd. London.
- Patel, R. M. 1962. Selection among factorially classified variables. Unpublished Ph.D. Thesis, North Carolina State College, Raleigh. (University Microfilms, Inc., Ann Arbor, Michigan).
- Rao, C. R. 1947. General methods of analysis for incomplete block designs. *J. Am. Stat. Assoc.* 42:541-561.
- Robinson, H. F., Rigney, J. A., and Harvey, P. H. 1948. Investigations in plot technique with peanuts. *N. C. Agr. Exp. Sta. Tech. Bul.* 86.
- Smith, H. F. 1936. A discriminant function for plant selection. *Ann. Eugenics* 7:240-250.
- Smith, H. F. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. *J. Agr. Sci.* 28:1-23.
- Yates, F. 1936. A new method of arranging variety trials involving a large number of varieties. *J. Agr. Sci.* 26:424-455.

INSTITUTE OF STATISTICS
NORTH CAROLINA STATE COLLEGE

(Mimeo Series available for distribution at cost)

283. Schutzenberger, M. P. On the recurrence of patterns. April, 1961.
284. Bose, R. C. and I. M. Chakravarti. A coding problem arising in the transmission of numerical data. April, 1961.
285. Patel, M. S. Investigations on factorial designs. May, 1961.
286. Bishir, J. W. Two problems in the theory of stochastic branching processes. May, 1961.
287. Konsler, T. R. A quantitative analysis of the growth and regrowth of a forage crop. May, 1961.
288. Zaki, R. M. and R. L. Anderson. Applications of linear programming techniques to some problems of production planning over time. May, 1961.
289. Schutzenberger, M. P. A remark on finite transducers. June, 1961.
290. Schutzenberger, M. P. On the equation $a^{2+n} = b^{2+m}c^{2+p}$ in a free group. June, 1961.
291. Schutzenberger, M. P. On a special class of recurrent events. June, 1961.
292. Bhattacharya, P. K. Some properties of the least square estimator in regression analysis when the 'independent' variables are stochastic. June, 1961.
293. Murthy, V. K. On the general renewal process. June, 1961.
294. Ray-Chaudhuri, D. K. Application of geometry of quadrics of constructing PBIB designs. June, 1961.
295. Bose, R. C. Ternary error correcting codes and fractionally replicated designs. May, 1961.
296. Koop, J. C. Contributions to the general theory of sampling finite populations without replacement and with unequal probabilities. September, 1961.
297. Foradori, G. T. Some non-response sampling theory for two stage designs. Ph.D. Thesis. November, 1961.
298. Mallios, W. S. Some aspects of linear regression systems. Ph.D. Thesis. November, 1961.
299. Taeuber, R. C. On sampling with replacement: an axiomatic approach. Ph.D. Thesis. November, 1961.
300. Gross, A. J. On the construction of burst error correcting codes. August, 1961.
301. Srivastava, J. N. Contribution to the construction and analysis of designs. August, 1961.
302. Hoeffding, Wassily. The strong laws of large numbers for u-statistics. August, 1961.
303. Roy, S. N. Some recent results in normal multivariate confidence bounds. August, 1961.
304. Roy, S. N. Some remarks on normal multivariate analysis of variance. August, 1961.
305. Smith, W. L. A necessary and sufficient condition for the convergence of the renewal density. August, 1961.
306. Smith, W. L. A note on characteristic functions which vanish identically in an interval. September, 1961.
307. Fukushima, Kozo. A comparison of sequential tests for the Poisson parameter. September, 1961.
308. Hall, W. J. Some sequential analogs of Stein's two-stage test. September, 1961.
309. Bhattacharya, P. K. Use of concomitant measurements in the design and analysis of experiments. November, 1961.
310. Anderson, R. L. Designs for estimating variance components. November, 1961.
311. Guzman, Miguel Angel. Study and application of a non-linear model for the nutritional evaluation of proteins. Ph.D. Thesis—January, 1962.
312. Koop, John C. An upper limit to the difference in bias between two ratio estimates. January, 1962.
313. Prairie, Richard R. and R. L. Anderson. Optimal designs to estimate variance components and to reduce product variability for nested classifications. January, 1962.
314. Eicker, Friedhelm. Lectures on time series. January, 1962.
315. Potthoff, Richard F. Use of the Wilcoxon statistic for a generalized Behrens-Fisher problem. February, 1962.
316. Potthoff, Richard F. Comparing the medians of the symmetrical populations. February, 1962.
317. Patel, R. M., C. C. Cockerham and J. O. Rawlings. Selection among factorially classified variables. February, 1962.
318. Amster, Sigmund. A modified Bayes stopping rule. April, 1962.
319. Novick, Melvin R. A Bayesian indifference postulate. April, 1962.
320. Potthoff, Richard F. Illustration of a technique which tests whether two regression lines are parallel when the variances are unequal. April, 1962.
321. Thomas, R. E. Preemptive disciplines for queues and stores. April, 1962. Ph.D. Thesis.
322. Smith, W. L. On the elementary renewal theorem for non-identically distributed variables. April, 1962.
323. Potthoff, R. F. Illustration of a test which compares two parallel regression lines when the variances are unequal. May, 1962.
324. Bush, Norman. Estimating variance components in a multi-way classification. Ph.D. Thesis. June, 1962.
325. Sathe, Yashwant Sedashiv. Studies of some problems in non-parametric inference. June, 1962.
326. Hoeffding, Wassily. Probability inequalities for sums of bounded random variables. June, 1962.
327. Adams, John W. Autoregressive models and testing of hypotheses associated with these models. June, 1962.