# AUGMENTED (OR HOONUIAKU) DESIGNS[1/]

Walter T. Federer[2/]

BU-74-M

February, 1956

---

1/ This paper is a result of a discussion among F. C. Denison, E. B. Holroyde, J. F. Morgan, Jr., A. J. Mangelsdorf, T. Nakayama, and the author at Kahuku Plantation on March 8, 1955. An augmented randomized complete block design was installed at Kahuku on April 11, 1955. Additional experiments with this design have been installed in 1955 on the Kahuku, Ewa, and Pioneer Mill Plantations.

When these designs were first developed they were named "chain block designs" and this is the title used in the first manuscripts and lectures on the designs. Since these designs are different from the ordinary chain block and linked block designs, a new name should be applied to them. Since a name should be descriptive and popular, a number of people were consulted. The author wishes to express thanks to the various people offering names. Some of the people offering suggestions and some of the suggested names are: W. S. Conner (designs with tagalongs), J. W. Tukey (caboose, penthouse, piggy-back, supplemental), O. Kempthorne (augmented), W. G. Cochran (augmented), and associates at Cornell University (bonded, fettered, expanded, ranged, topped, adjoined, "69", "combonded", bolted, latched, entangled, raveled, conjoined, etc.). From all names considered, it appears that "augmented" is most descriptive. Another appropriate name for these designs might be "hoonuiaku" (pronounced hō•ō•nōō•ee•ä•kōō) since the designs were developed in Hawaii.

2/ Principal Statistician, Experiment Station, HSPA and Consultant in Statistics, PRI, while on sabbatical leave from Cornell University from September, 1954 to September, 1955.

---

## INTRODUCTION

A number of chain block and linked block designs have been described by Youden *et al*. [1,4,7,8]. These designs were developed to meet the needs of physicists and chemists whose experimental material is relatively homogeneous, thus allowing reliable estimates with only one or two measurements. On the other hand, field results for breeding experiments are quite variable, necessitating several measurements over time and space. However, in the early stages of a breeding program, seed is often the limiting factor and it is impossible to plant more than one or two plots of the material under consideration (seedlings). It would appear, then, that some design similar to a chain block design is indicated because of the scarcity rather than the homogeneity of the experimental material. Sufficient material of standard and promising new varieties is available to plant several replicates. If the varieties and seedlings are grown together in the same experimental design, comparisons among varieties, among seedlings, and among varieties and seedlings are possible.

At present, the seedlings are grown in 30' x 30' single plots with a check variety every third plot [3,5,6] and the varieties are grown in 40' x 40' plots in replicated tests. Since the use of a check variety in every third plot is not an efficient design [3] and since no experimental error is available for comparing the yields of one seedling with another, it was desired to have an experimental design which would afford comparisons of varieties with seedlings and to have a measure of the experimental error for seedlings.

The purpose of this paper is to present some new designs which will achieve the following objectives:

(i) provide efficient designs for seedlings,

(ii) provide a measure of the experimental error for seedlings,

(iii) provide comparisons of seedlings with varieties and varieties with varieties, and

(iv) combine seedling tests with Grade A variety tests.

The plot size for these experiments will be 40' x 40' with 10' of border sugar cane being discarded on the sides and ends of the plot. This is the plot size established by the late R. J. Borden and others as the plot size for Grade A variety tests.

A number of types of a new class of designs, augmented designs, are presented. The solutions for two of the new designs are given with examples. The solution for the other types may be obtained in the same manner. However, the algebra will be somewhat more complicated.

A discussion of the construction of chain block designs is included because an attempt was made to use these designs for the experimental situation encountered and because a study of these designs led to the development of augmented designs. Notation used is similar to that used by Youden and Connor [8].

The success of any experimental design depends upon the ability of the experimenter to stratify or block the area into blocks which are relatively homogeneous within each block. For a large number of designs, the block size is constant. For chain block, linked block, and augmented designs, the block size may vary, but the blocks should be set up to minimize the variation among plots within the blocks.

CHAIN BLOCK DESIGNS

In the chain block design proposed by Youden and Connor [8], there are basically two replicates for $v_2$ treatments (varieties) and one replicate on $v_1$ treatments (seedlings). The value of $v_2$ and the number, $2n_2$, of varieties included in an incomplete block determine the number of incomplete blocks, b; thus, $b = v_2/n_2$. To illustrate, suppose $v_2 = 3(A,B,C)$, r (no. of replicates) = 2, and $n_2 = 1$; then, b = 3. The grouping of the varieties in a block is

| Group 1 | Group 2 | Group 3 |
|---|---|---|
| A | C | B |
| B | A | C |
| $\frac{n_1}{3}$ seedlings | $\frac{n_1}{3}$ or $\frac{n_1}{3}$ + 1 seedlings | $\frac{n_1}{3}$ or $\frac{n_1}{3}$ + 1 seedlings |

Thus, Group 1 is "chained" or "linked" to Group 2 through variety A, Group 2 is linked to Group 3 through variety C, and Group 3 is linked to Group 1 through variety B. The number of seedlings, $n_{1j}$, included in each block is determined only by the number of seedlings available and the system of blocking used. In the examples cited here an attempt was made to equalize the number of seedlings in each block.

As a further example, consider the chain block designs available for $v_2 = 6$:

Group or Block

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| A | B | C | D | E | F |
| F | A | B | C | D | E |
| $n_{11}$ | $n_{12}$ | $n_{13}$ | $n_{14}$ | $n_{15}$ | $n_{16}$ |

$v_2 = 6$, $n_2 = 1$

$b = 6$, $r = 2$

$n_{1j} = (n_{1/6}) + 1$ or $n_{1/6}$ to the largest whole integer.

Group

| 1 | 2 | 3 |
|---|---|---|
| A | E | C |
| B | F | D |
| C | A | E |
| D | B | F |
| $n_{11}$ | $n_{12}$ | $n_{13}$ |

$v_2 = 6$, $n_2 = 2$

$b = 3$, $r = 2$

$n_{1j} = (n_{1/3}) + 1$ or $n_{1/3}$ to the largest whole integer.

Chain block designs for values of $v_2$ up to 25 are listed in Table 1.

TABLE 1. CHAIN BLOCK DESIGNS FOR TWO REPLICATES FOR VARIOUS VALUES OF $v_2$

| $v_2$ | $n_2$ | b | $df_e$ | $v_2$ | $n_2$ | b | $df_e$ |
|---|---|---|---|---|---|---|---|
| 3 | 1 | 3 | 1 | 16 | 4 | 4 | 13 |
| 4 | 1 | 4 | 1 | 17 | 1 | 17 | 1 |
| 5 | 1 | 5 | 1 | 18 | 1 | 18 | 1 |
| 6 | 1 | 6 | 1 | 18 | 2 | 9 | 10 |
| 6 | 2 | 3 | 4 | 18 | 3 | 6 | 13 |
| 7 | 1 | 7 | 1 | 18 | 6 | 3 | 16 |
| 8 | 1 | 8 | 1 | 19 | 1 | 19 | 1 |
| 8 | 2 | 4 | 5 | 20 | 1 | 20 | 1 |
| 9 | 1 | 9 | 1 | 20 | 2 | 10 | 11 |
| 9 | 3 | 3 | 7 | 20 | 4 | 5 | 16 |
| 10 | 1 | 10 | 1 | 20 | 5 | 4 | 17 |
| 10 | 2 | 5 | 6 | 21 | 1 | 21 | 1 |
| 11 | 1 | 11 | 1 | 21 | 3 | 7 | 15 |
| 12 | 1 | 12 | 1 | 21 | 7 | 3 | 19 |
| 12 | 2 | 6 | 7 | 22 | 1 | 22 | 1 |
| 12 | 3 | 4 | 9 | 22 | 2 | 11 | 12 |
| 12 | 4 | 3 | 10 | 23 | 1 | 23 | 1 |
| 13 | 1 | 13 | 1 | 24 | 1 | 24 | 1 |
| 14 | 1 | 14 | 1 | 24 | 2 | 12 | 13 |
| 14 | 2 | 7 | 8 | 24 | 3 | 8 | 17 |
| 15 | 1 | 15 | 1 | 24 | 4 | 6 | 19 |
| 15 | 3 | 5 | 11 | 24 | 6 | 4 | 21 |
| 15 | 5 | 3 | 13 | 24 | 8 | 3 | 22 |
| 16 | 1 | 16 | 1 | 25 | 1 | 25 | 1 |
| 16 | 2 | 8 | 9 | 25 | 5 | 5 | 21 |

The number of degrees of freedom for the error mean square ($df_e$) is determined by the values of $v_2$, $n_2$, and b. The following analysis of variance table illustrates the partitioning of the degrees of freedom:

| Source of Variation | Degrees of Freedom |
|---|---|
| Among blocks (groups) | $b - 1$ |
| Among entries | $v - 1$ |
| Error mean square | $v_2 - b + 1 = df_e$ |

where $v_1 + v_2 = v$, $n_2 b = v_2$, the total number of plots $= N = v_1 + 2v_2$, and $v_2 - b + 1 = v_2(1 - \frac{1}{n_2}) + 1$.

Thus, the degrees of freedom for the above three designs are 1, 1, and 4, respectively. The degrees of freedom for the remaining designs in Table 1 are determined in the same manner. Since it is desired to have around 10 to 20 degrees of freedom associated with the error mean square, the designs with only one degree of freedom would seldom, if ever, be used.

One way of increasing $df_e$ is to increase the number of replicates. This may be accomplished by using the generalized chain block design as described by Mandel [4]. For example, consider the following design, omitting the $v_1$ treatments which are replicated once:

| Group | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| A | G | E | C |
| B | H | F | D |
| C | A | G | E |
| D | B | H | F |
| E | C | A | G |
| F | D | B | H |

$v_3 = 8,\ r = 3$

$n_2 = 4,\ b = 4$

$df_e = 13$

Group 1 is chained to group 2 by varieties A, B, C, D. Group 2 is chained to group 3 by varieties G, H, A, B; etc. This process may be used to set up chain block designs with more than two replicates. Mandel [4] also describes chain block designs for two-way control of heterogeneity.

## AUGMENTED (OR HOONUIAKU) DESIGNS

The class of experimental designs discussed in the present section is different from the class known as chain block designs in that the latter class has a common link to only one other block, whereas the former class has several common links. However, it may be more enlightening to consider an augmented design as a standard design plus additional treatments (the seedlings) in the blocks or cells of the design rather than as a design which has multiple links. The error mean square from all augmented designs with the varieties repeated b times and the seedlings in once is obtained from the analysis of the data on the varieties only. Hence, from an analysis standpoint, the consideration of a standard design with additional entries in the blocks or cells is more appropriate than the consideration of multiple linking.

As will be evident from the following, an augmented design may be formed from any standard design simply by enlarging the number of experimental units in the complete block, the incomplete block, the cell, etc., in the basic standard design. The adjustments for the $v_1$ seedling totals depend upon the design used and their location in the design. The totals of the entries repeated b times, the varieties, may or may not require adjustment. For example, no adjustments are required in the augmented randomized complete block design, but adjustments are required in the augmented triple lattice design.

## I. Augmented Completely Randomized Design

The first experimental design that comes to mind for $v_b$ varieties repeated b times and for $v_1$ seedlings each in one plot is the completely randomized design [2]. (The subscript on v denotes the number of times an entry is replicated.) The total experimental area is divided up into $bv_b + v_1 = N$ plots and the varieties and seedlings are allotted to the plots completely at random. No blocking is made of the entire area. By chance, all variety plots could fall in one corner of the experimental area. The analysis of variance for this well-known experimental design is:

| Source of Variation | Degrees of Freedom |
|---|---|
| Entries | $v_b + v_1 - 1$ |
| Among varieties | $v_b - 1$ |
| Among seedlings | $v_1 - 1$ |
| Varieties vs. seedlings | 1 |
| Within entries | $N - v_b - v_1 = v_b(b-1)$ |
| Total | $N - 1$ |

The variety and seedling means do not require adjustment. The standard error of a mean difference for two varieties is

$$\sqrt{\text{error mean square } (2/b)} \qquad \text{(I-1)}$$

The standard error of a difference between two seedling means is

$$\sqrt{2 \text{ (error mean square)}}. \qquad \text{(I-2)}$$

The standard error of a difference between a variety and a seedling mean is

$$\sqrt{\text{error mean square } (1 + 1/b)} \qquad \text{(I-3)}$$

This design is not recommended for sugar cane or other field experiments because it is highly improbable that a uniform area of N plots could be found. In most cases, it will be possible to reduce the error mean square considerably by proper stratification of the experimental area.

## II. Augmented Randomized Complete Block Design

An augmented design which holds considerable promise for sugar cane variety and seedling tests is the one in which the $v_b$ varieties are included once in each of the b blocks (i.e., replicated b times) and $v_1$ seedlings are included once in one of the b blocks. The blocks contain the $v_b$ varieties plus $n_{1j}$ seedlings, or a total of $v_b + n_{1j} = N_j$ plots.

$$bv_b + \sum_{j=1}^{b} n_{1j} = bv_b + v_1 = \sum_{j=1}^{b} N_j = N. \qquad \text{(II-1)}$$

The total number of replicates on the $v_b + v_1 = v$ entries is

$$\frac{bv_b + v_1}{v_b + v_1} = 1 + \frac{(b-1)\ v_b}{v} . \qquad \text{(II-2)}$$

To illustrate the grouping of varieties and seedlings, consider the following design for $v_b = 4$ (A,B,C,D), $v_1 = 11$ (e,f,g,h,i,j,k,l,m,n,o), and $b = 4$:

| Group 1 | Group 2 | Group 3 | Group 4 |
|---|---|---|---|
| A | A | A | A |
| B | B | B | B |
| C | C | C | C |
| D | D | D | D |
| e | h | k | n |
| f | i | l | o |
| g | j | m | |

Thus, there will be seven plots in three of the blocks and six plots in the fourth block. As is apparent from the above, the design is completely flexible with regard to the values of $v_b$, b, and $v_1$.

### Randomization

As indicated earlier, there are b blocks with $N_j$ plots in the j'th block. The randomization procedure follows:

(i) Randomly allot the varieties to the $v_b + n_{11} = N_1$ plots in block 1. This process is continued for the remaining blocks using a new random allotment in each block.

(ii) Randomly allot the $v_1$ seedlings to the remaining plots.

In setting up the b blocks, care should be taken to include a relatively homogeneous area within each block. The purpose of the blocking is to remove as much as possible of the heterogeneity in the experimental area and to have relatively homogeneous plots within each block.

Analysis

The analysis of the results for this experiment differs from that for orthogonal designs like the randomized complete block and latin square designs. The variety means, $\bar{y}_i$, require no adjustment for blocks since all varieties appear in all blocks (see formula (A19)). The mean effect, m, in the experiment is estimated as (see formula (A17) and (A18)):

$$m = \frac{1}{v_b + v_1} (Y_{...} - (b-1)\Sigma\bar{y}_{i.} - \Sigma n_{1j} r_j). \qquad \text{(II-3)}$$

If $n_{1j}$ is a constant then

$$m = \frac{1}{v_b + v_1} (Y_{...} - (b-1)\Sigma\bar{y}_{i.}). \qquad \text{(II-4)}$$

The block effect is obtained from the following:

$$r_j = \frac{1}{v_b}(Y_{.j.} - \Sigma\bar{y}_{i.} - \sum_{g=1}^{n_{1j}} Y_{1jg}). \qquad \text{(II-5)}$$

The adjusted yield for a seedling plot is estimated from the following equation:

$$m + t_{1jg} = Y_{1jg}' = Y_{1jg} - r_j. \qquad \text{(II-6)}$$

The symbols in the above equations are defined in Appendix A.

The analysis of variance for an augmented randomized complete block design is:

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Square |
|---|---|---|---|
| Block (ignoring entry) | $b - 1$ | B | -- |
| Entries = varieties and sdlgs. (eliminating blocks) | $v_b+v_1-1$ | $SS_{vs}$ | $V_{vs}$ |
| Varieties | $v_b-1$ | $SS_v$ | $V_v$ |
| Sdlgs. and varieties vs. sdlgs. | $v_1$ | $SS_s$ | $V_s$ |
| Intrablock error | $(v_b-1)(b-1)$ | $SS_e$ | $E_e$ |
| Total | $N - 1$ | $SS_t$ | -- |

An alternative analysis for the above design is:

| Source of Variation | Degrees of Freedom | Sum of Squares | Mean Squares |
|---|---|---|---|
| Block (eliminating varieties and seedlings) | $b - 1$ | $SS_b$ | $E_b$ |
| Entry (ignoring block) | $v_b + v_1 - 1$ | $T$ | -- |
| Intrablock error | $(v_b - 1)(b - 1)$ | $SS_e$ | $E_e$ |
| Total | $N - 1$ | $SS_t$ | -- |

The various sums of squares are computed as follows (see Appendix A):

$$SS_t = \sum_{j=1}^{b} \sum_{i=1}^{v} Y_{bij}^2 + \sum_{j=1}^{b} \sum_{g=1}^{n_{1j}} Y_{1jg}^2 - \frac{Y_{\cdot\cdot\cdot}^2}{N} \tag{II-7}$$

$$B = \Sigma \frac{Y_{\cdot j \cdot}^2}{v_b + n_{1j}} - \frac{Y_{\cdot\cdot\cdot}^2}{N} . \tag{II-8}$$

$SS_{vs}$ = Sum of squares due to m, $r_j$, $t_{bi}$, and $t_{1jg}$ minus sum of squares due to $m^*$, and $r_j^*$ (assuming each $t_{bi}$ and $t_{1jg} = 0$) $= mY_{\cdot\cdot\cdot} + \sum_{j=1}^{b} r_j Y_{\cdot j \cdot}$

$$+ \sum_{i=1}^{v_b} t_{bi} Y_{i\cdot} + \sum_{j=1}^{b} \sum_{g=1}^{n_{1j}} t_{1jg} Y_{1jg} - \frac{\Sigma Y_{\cdot j \cdot}}{v_b + n_{1j}} . \tag{II-9}$$

$$SS_v = \frac{\Sigma Y_{i\cdot}^2}{b} - \frac{(\Sigma Y_{i\cdot})^2}{bv_b} \tag{II-10}$$

$$SS_s = SS_{vs} - SS_v \tag{II-11}$$

$$SS_e = SS_t - SS_{vs} - B \tag{II-12}$$

$$T = \frac{\Sigma Y_{i\cdot}^2}{b} + \underset{jg}{\Sigma\Sigma} Y_{1jg}^2 - \frac{Y_{\cdot\cdot\cdot}^2}{N} . \tag{II-13}$$

$$SS_b = B - (T - SS_{vs}). \tag{II-14}$$

The estimated variances of differences between various combinations are given in Appendix A.

An Example

To illustrate the analysis of an augmented randomized complete block design, a portion of the data from a uniformity (blank) trial on Field 78 at Pioneer Mill Sugar Company, 1931, was used to construct an example (Table 2). The three blocks are from three different level ditches, and the yield character is tons of sugar cane per acre (TCA). Four hypothetical varieties, A, B, C, and D, and eight hypothetical seedlings, e, f, g, h, i, j, k, and l are used. Blocks 1 and 3 contain $v_b + n_{11} = v_b + N_{13} = 4 + 3 = 7$ entries each, and block 2 contains $v_b + n_{12} = 4 + 2 = 6$ entries. The block and grand totals are given in Table 2, and the variety and seedling totals are given in Table 3. As stated above, the variety means do not require adjustment for block effects. Hence, the arithmetic or unadjusted means for the varieties are used.

The mean effect, m, and the block effects, $r_j$, are estimated as follows (see formulae (II-3) and (II-5)).

$$m = \frac{1}{4 \cdot 8}\left\{ 1630 - 2(329) - \frac{1}{4}\left\{3(-13) + 2(3) + 3(10)\right\}\right\} = 81.0625.$$

$$r_1 = \frac{1}{4}\left\{ 535 - 329 - 74 - 70 - 75 \right\} = \frac{-13}{4} = -3.25.$$

$$r_2 = \frac{3}{4} = 0.75$$

$$r_3 = \frac{10}{4} = 2.50.$$

The sum of the block effects should add to zero, thus $-3.25 + 0.75 + 2.50 =$ zero.

The variety effects are computed from formula (A19). For example, the effect for variety A is $\bar{y}_{A\circ} - m = 84.6667 - 81.0625 = 3.6042$. The remaining variety effects are similarly computed and are presented in Table 3.

The seedling effects are computed from formula (II-6). For example, the effect for seedling l is $Y_{111} - r_1 - m = 74 - 81.0625 - (-3.25) = -3.8125$. The remaining seedling effects are presented in Table 3. As a partial check, the sum of the variety and of seedling effects should equal zero (formula (A16)).

The computations for the sums of squares in the analysis of variance (Table 4) are presented below (see formulae (II-7) to (II-14)):

$$SS_t = 74^2 + 78^2 + \ldots + 79^2 + 82^2 - \frac{1630^2}{20}$$

$$= 133,652 - 132,845 = 807.$$

TABLE 2. FIELD ARRANGEMENT OF YIELDS (TCA) FOR AN AUGMENTED RANDOMIZED COMPLETE BLOCK DESIGN. DATA FROM BLANK TEST - PIONEER MILL #78-1931

| | Block 1 | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| Variety or seedling | l | C | D | g | A | B | k | |
| Yield | 74 | 78 | 78 | 70 | 83 | 77 | 75 | 535 = $Y_{.1.}$ |
| | Block 2 | | | | | | | Total |
| Variety or seedling | D | B | A | C | e | i | - | |
| Yield | 91 | 81 | 79 | 81 | 79 | 78 | - | 489 = $Y_{.2.}$ |
| | Block 3 | | | | | | | Total |
| Variety or seedling | h | C | A | f | D | B | j | |
| Yield | 96 | 87 | 92 | 89 | 81 | 79 | 82 | 606 = $Y_{.3.}$ |
| | | | | | | | | 1630 = $Y_{...}$ |

TABLE 3. VARIETY AND SEEDLING TOTALS, EFFECTS, AND ADJUSTED MEANS

| Variety | Total | Means | Effects |
|---|---|---|---|
| A | 254 | 84.67 | 3.6042 |
| B | 237 | 79.00 | -2.0625 |
| C | 246 | 82.00 | 0.9375 |
| D | 250 | 83.33 | 2.2708 |
| Total | 987 | 329.00 | 4.7500 |

$m = 81.0625$
$r_1 = -3.25$
$r_2 = 0.75$
$r_3 = 2.50$

| Seedling | Unadj. Mean | Adjusted Mean | Effects, $t_{1jg}$ |
|---|---|---|---|
| e | 79 | 78.25 | -2.8125 |
| f | 89 | 86.50 | 5.4375 |
| g | 70 | 73.25 | -7.8125 |
| h | 96 | 93.50 | 12.4375 |
| i | 78 | 77.25 | -3.8125 |
| j | 82 | 79.50 | -1.5625 |
| k | 75 | 78.25 | -2.8125 |
| l | 74 | 77.25 | -3.8125 |
| Total | 643 | 643.75 | -4.7500 |

$$B = \frac{535^2 + 606^2}{7} + \frac{489^2}{6} - \frac{1630^2}{20}$$

$$= 133{,}205.0714 - 132{,}845 = 360.0714$$

$$SS_{vs} = 81.0625(1630) + [(-3.25)(535) + .75(489)$$

$$+ 2.50(606)] + [3.6042(254) + \ldots + 2.2708(250)]$$

$$+ [79(-2.8125) + 89(5.4375) + \ldots + 74(-3.8125)]$$

$$- 133{,}205.0714 = 133{,}490.1668 - 133{,}205.0714$$

$$= 285.0954.$$

$$SS_v \frac{254^2 + \ldots + 250^2}{3} - \frac{987^2}{12} = 52.9167.$$

$$SS_s = 285.0954 - 52.9167 = 232.1787$$

$$SS_e = 807 - 360.0714 - 285.0954 = 161.8332$$

$$T = \frac{254^2 + \ldots + 250^2}{3} + 74^2 + \ldots + 75^2$$

$$- \frac{1630^2}{20} = 575.6667.$$

$$SS_b = 360.0714 - (575.6667 - 285.0954) = 69.5001.$$

If it is desired to obtain only an estimate of the error variance and to use one of the multitple range tests [see 2, Chapter II], a simple analysis is available. This analysis consists of using only the data from the varieties (Table 5) and using the analysis for a randomized complete block design for an orthogonal array. The fact that $SS_b$ is equal to the blocks sum of squares in Table 6 within rounding errors, is not a coincidence. These sums of squares must be equal. Likewise, the error sum of squares in Table 6 must equal SS, within rounding errors. This feature is a property of the augmented designs described herein.

The various standard errors of a mean difference are computed as follows:

Between two variety means

$$\sqrt{\frac{2}{3}\ (26.9722)} = 4.24.$$

Between two seedlings in the same block

$$\sqrt{2\ (26.9722)} = 7.34.$$

Between two seedlings not in the same block

$$\sqrt{2\ (26.9722)(1 + \frac{1}{4})} = 8.21.$$

Between a variety and a seedling

$$\sqrt{26.9722(1 + \frac{1}{3} + \frac{1}{4} + \frac{1}{12})} = 6.70$$

TABLE 4. ANALYSES OF VARIANCE

| Source of Variation | d.f. | | Sum of Squares | Mean Square |
|---|---|---|---|---|
| Blocks (ignoring treatment) | 2 | | 360.0714 | 180.04 |
| Treatments (eliminating blocks) | 11 | | 285.0954 | 25.92 |
| Varieties | | 3 | 52.9167 | 17.64 |
| Seedlings and varieties vs. seedlings | | 8 | 232.1787 | 29.02 |
| Error | 6 | | 161.8332 | 26.9722 |
| Total | 19 | | 807.0000 | -- |
| Blocks (eliminating treatments) | 2 | | 69.5001 | |
| Treatments(ignoring blocks) | 11 | | 575.667 | -- |
| Varieties | | 3 | 52.9167 | -- |
| Seedlings | | 7 | 505.8750 | -- |
| Seedlings vs. varieties | | 1 | 16.8750 | -- |
| Error | 6 | | 161.8332 | -- |

TABLE 5. YIELD DATA FOR VARIETIES

| | Block | | | |
|---|---|---|---|---|
| Variety | 1 | 2 | 3 | Total |
| A | 83 | 79 | 92 | 254 |
| B | 77 | 81 | 79 | 237 |
| C | 78 | 81 | 87 | 246 |
| D | 78 | 91 | 81 | 250 |
| Total | 316 | 332 | 339 | 987 |

TABLE 6. ANALYSIS OF VARIANCE FOR DATA IN TABLE 5

| Source of Variation | d.f. | S.S. | M.S. |
|---|---|---|---|
| Blocks | 2 | 69.5000 | 34.75 |
| Varieties | 3 | 52.9167 | 17.64 |
| Error | 6 | 161.8333 | 26.9722 |
| Total | 11 | 284.2500 | -- |

For Tukey's hsd test [see 2, Chapter II], the first three standard errors of a mean difference above are multiplied by $q_{05}$ (for 6 d.f. and 12 treatments) divided by $\sqrt{2}$. For the last standard error, it is suggested that the $q_{05}$ value be multiplied by $\sqrt{E_e/n'}$ where $1/n'$ = harmonic mean of the coefficients for the standard errors of a difference. Thus, $n' = 4/(1 + 1/3 + 1/4 + 1/12) = 12/5$. It is not known how appropriate this approximation is for means with unequal numbers of observations.

## III. Augmented Latin Square Design

In this design, the $v_b = b$ varieties are arranged in such a way that each variety appears once in a row or once in a column. This makes use of the principles of the latin square design. For example, consider the following design with $v_b = b = 3$ varieties (A, B, and C) in b = 3 rows and columns with $v_1 = 16$ seedlings (1,2,...,16, where the numbers are randomly allotted to the seedlings):

| | Column 1 | | | Column 2 | | | Column 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Row 1 | C | 1 | 2 | A | 3 | 4 | 5 | B | 6 |
| 2 | A | 12 | 11 | 10 | B | 9 | 8 | C | 7 |
| 3 | | B | 13 | 14 | 15 | C | 16 | A | |

The number of seedlings in a given row and column, $n_{1hi}$, need not be the same; this is illustrated in the above design. The analysis will be somewhat simpler if $n_{1hi}$ is a constant. Also, the $n_{1hi}$ should be made as nearly equal as possible subject to the heterogeneity in the experimental area. It should be noted that the rows (or columns) could be laid end to end if desired [see 2].

### Randomization

The randomization scheme for the augmented latin square design follows:

(i) Select a random arrangement for the $v_b = b$ varieties in a bxb latin square design [see 2, Chapter VI].

(ii) There are $n_{1hi} + 1 = N_{hi}$ plots in the i'th row and j'th column. The variety falling in the h'th row and i'th column is randomly allocated to one of the $N_{hi}$ plots.

(iii) Assign random numbers to the $v_1$ seedlings and then distribute these to the remaining plots.

Analysis

The analysis of the data for an augmented latin square design follows that for an augmented randomized complete block design fairly closely. The variety means do not require adjustment; the seedling means must be adjusted for a row effect as well as a column effect. These adjustments are described in Appendix B and in the example.

The following analysis of variance is performed on the varieties only:

| Source of Variation | d.f. | Sum of Squares | Mean Square |
|---|---|---|---|
| Row | b-1 | $SS_r$ | - |
| Column | b-1 | $SS_c$ | - |
| Variety | b-1 | $SS_v$ | - |
| Residual or error | (b-1)(b-2) | $SS_e$ | $E_e$ |
| Total | $b^2$-1 | -- | - |

The analysis of variance for varieties and seedlings is:

| Source of Variation | d.f. | | Sum of Squares | |
|---|---|---|---|---|
| Row (ignoring entry and column) | b-1 | | R | |
| Column (ignoring entry; eliminating row | b-1 | | C | |
| Entry (eliminating row and column) | $v_1$+ b-1 | | $SS_{vs}$ | |
| Variety | | b-1 | | $SS_v$ |
| Seedling and seedling vs. variety | | $v_1$ | | $SS_s$ |
| Error | (b-1)(b-2) | | $SS_e$ | |
| Total | $b^2 + v_1 - 1$ | | - | |

The sum of squares for $SS_{vs}$ may be obtained from the following relation:

Sum of squares for varieties and seedlings (ignoring row and column effects)

$$- (R - SS_r + C - SS_c) = SS_{vs}.$$

The various standard errors are described in Appendix B.

## An Example

A hypothetical numerical example was constructed to illustrate the analysis for an augmented latin square design. Given that $m = 10$, $r_1 = -1$, $r_2 = 0$, $r_3 = 1$, $c_1 = -3$, $c_2 = -1$, $c_3 = 4$, $t_A = -1$, $t_B = -2$, $t_C = -3$, $t_d = 0$, $t_e = 2$, and $t_f = 4$, the following example was constructed:

**Yields For Augmented Latin Square Design**

| Rows | Columns 1 | | 2 | 3 | | Totals: All Entries | Totals: Varieties Only |
|---|---|---|---|---|---|---|---|
| 1 | | (A) 5 | (B) 6 | (C) 10 | (d) 13 | $34 = Y_{\cdot 1 \cdot\cdot}$ | $21 = Y_{b \cdot 1 \cdot}$ |
| 2 | | (B) 5 | (C) 6 | (e) 16 | (A) 13 | $40 = Y_{\cdot 2 \cdot\cdot}$ | $24 = Y_{b \cdot 2 \cdot}$ |
| 3 | (C) 5 | (f) 12 | (A) 9 | (B) 13 | | $39 = Y_{\cdot 3 \cdot\cdot}$ | $27 = Y_{b \cdot 3 \cdot}$ |
| Totals (all) | $27 = Y_{\cdot\cdot 1 \cdot}$ | | $21 = Y_{\cdot\cdot 2 \cdot}$ | $65 = Y_{\cdot\cdot 3 \cdot}$ | | $113 = Y_{\cdot\cdot\cdot\cdot}$ | -- |
| Totals (var) | $15 = Y_{b \cdot 1 \cdot}$ | | $21 = Y_{b \cdot 2 \cdot}$ | $36 = Y_{b \cdot 3 \cdot}$ | | $72 = Y_{b \cdot\cdot\cdot}$ | -- |

The variety and seedling (unadjusted) totals and means are:

| | |
|---|---|
| $Y_{b \cdot\cdot A} = 27$ | $\bar{y}_{b \cdot\cdot A} = 9$ |
| $Y_{b \cdot\cdot B} = 24$ | $\bar{y}_{b \cdot\cdot B} = 8$ |
| $Y_{b \cdot\cdot C} = 21$ | $\bar{y}_{b \cdot\cdot C} = 7$ |
| $Y_{1 \cdot\cdot d} = 13$ | -- |
| $Y_{1 \cdot\cdot e} = 16$ | -- |
| $Y_{1 \cdot\cdot f} = 12$ | -- |

From formulae (B10) to (B14) we note that

$$m = \frac{1}{3+3}\left\{113 - \left(3 - 1 - \frac{2(3)}{3}\right)(9 + 8 + 7) - \frac{1}{3}((1)(21) + (1)(24) + (1)(27)) - \frac{1}{3}((1)(15) + (0)(21) + 2(36))\right\} = 10,$$

$$r_1 = \frac{1}{3}\left\{21 - 24\right\} = -1,$$

$$r_2 = \frac{1}{3}\left\{24 - 24\right\} = 0,$$

$$r_3 = \frac{1}{3}\left\{27 - 24\right\} = 1,$$

$$c_1 = \frac{1}{3}\left\{15 - 24\right\} = -3,$$
$$c_2 = \frac{1}{3}\left\{21 - 24\right\} = -1,$$
$$c_3 = \frac{1}{3}\left\{36 - 24\right\} = 4,$$
$$t_A = 9 - 10 = -1,$$
$$t_B = 8 - 10 = -2,$$
$$t_C = 7 - 10 = -3,$$
$$t_d = 13 - (-1) - (4) - 10 = 0,$$
$$t_e = 16 - (0) - (4) - 10 = 2, \text{ and}$$
$$t_f = 12 - (1) - (-3) - 10 = 4.$$

Thus, all effects agree with the original values. This is a necessary condition for the correctness of the equations.

Before computing the sum of squares, the estimates $m'$, $c_i'$, and $r_h'$ computed from formulae (B2) to (B4) with each $t_{bj}$ and each $t_{lhik}$ set equal to zero, are required. Using the relation in (B7) and (B8) the following normal equations for the constructed example are obtained:

$$12m' + c_1' + 2c_3' = Y.... = 113,$$
$$4(r_1' + m') + c_3' = 34,$$
$$4(r_2' + m') + c_3' = 40,$$
$$4(r_3' + m') + c_1' = 39,$$
$$4(c_1' + m') + r_3' = 27,$$
$$3(c_2' + m') = 21, \text{ and}$$
$$5(c_3' + m') - r_3' = 65.$$

Solutions for the unknowns in the above formulae plus formulae (B7) and (B8) are:

$$m' = 8.8918919,$$
$$r_1' = -1.4932432,$$
$$r_2' = .0067568,$$
$$r_3' = 1.4864865,$$
$$c_1' = -2.5135135,$$
$$c_2' = -1.8918919, \text{ and}$$
$$c_3' = 4.4054054.$$

The adjusted seedling means (totals) are (formula (B14)):

$$Y_{1\cdot\cdot d}' = t_{111d} + m = Y_{1\cdot\cdot d} - r_1 - c_3 \qquad \text{(III-1)}$$
$$= 0 + 10 = 13 - (-1) - 4 = 10,$$
$$Y_{1\cdot\cdot e}' = 2 + 10 = 16 - (0) - 4 = 12, \text{ and}$$
$$Y_{1\cdot\cdot f}' = 4 + 10 = 12 - 1 - (-3) = 14.$$

If only an estimate of the error mean square is desired then the analysis of variance in the central portion of Table 7 is calculated.* This analysis is identical to that for a standard kxk latin square. The analysis would be completed with the calculation of the various error variances given by formulae (B15) to (B19). Then, one of the multiple range tests [see 2, Ch. II] would be used to complete the comparison of means.

In some situations the analysis of variance given in the top part of Table 7 might be desired. The various sums of squares are computed as follows:

Total ss corrected for the mean (11 d.f.):

$$5^2 + 5^2 + \ldots\ 13^2 + 13^2 - \frac{113^2}{12} = 1235 - 1064.0833 = 170.9167.$$

Row sum of squares (ignoring column and entry effect)(2 d.f.):

$$R = \frac{34^2 + 40^2 + 39^2}{4} - \frac{113^2}{12} = 5.1667$$

Error sum of squares (2 d.f.):

$$\text{total S.S. uncorrected} - \Big\{ mY_{\cdot\cdot\cdot\cdot} + \Sigma r_h Y_{\cdot h\cdot\cdot} + \Sigma c_i Y_{\cdot\cdot i\cdot} + \Sigma t_{bj} Y_{b\cdot\cdot j} + \Sigma\Sigma\Sigma t_{1hik} Y_{1hik} \Big\} \qquad \text{(III-2)}$$

$$1235 - \Big\{ 10(113) + [-1(34) + 0(40) + 1(39)] + [-3(27) - 1(21) + 4(65)] + [-1(27) - 2(24) - 3(21)] + [0(13) + 2(16) + 4(12)] \Big\}$$
$$= 1235 - \{1130 + 5 + 158 - 138 + 80\}$$
$$= 0,$$

as it should for this example.

---

* In the example, the error sum of squares turns out to be zero as it should since no allowance was made for error variation in constructing the example.

TABLE 7. ANALYSES OF VARIANCE

All Yields

| Source of Variation | d.f. | Sum of Squares | Mean Square |
|---|---|---|---|
| Row (ignoring col. and entry) | 2 | 5.1667 | 2.5834 |
| Column (ignoring entry, eliminating row) | 2 | 121.7635 | 60.8818 |
| Entry (elim. row and column) | 5 | 43.9865 | 8.7973 |
| Variety | 2 | 6.0000 | 3.0000 |
| Remainder | 3 | 37.9865 | 12.6622 |
| Error | 2 | 0.0000 | 0.0000 |
| Total | 11 | 170.9167 | -- |
| Correction for mean | 1 | 1064.0833 | -- |
| Total (uncorrected) | 12 | 1235.0000 | -- |

On Varieties Yields Only

| | | | |
|---|---|---|---|
| Row | 2 | 6 | 3 |
| Column | 2 | 78 | 39 |
| Variety | 2 | 6 | 3 |
| Error | 2 | 0 | 0 |
| Total | 8 | 90 | -- |
| Correction for mean | 1 | 576 | -- |

On All Yields Assuming Zero Entry Effect

| | | | |
|---|---|---|---|
| Row(eliminating column) | 2 | 16.7635 | -- |
| Column (ignoring row) | 2 | 110.1667 | -- |
| Residual | 4 | 43.9865 | -- |
| Row (ignoring column) | 2 | 5.1667 | -- |
| Column (eliminating row) | 2 | 121.7635 | -- |
| Residual | 4 | 43.9865 | -- |

Entry (eliminating row and column effects) ss (5 d.f.):

$$SS_{vs} = mY_{....} + \Sigma r_h Y_{.h..} + \Sigma c_i Y_{..i.} + \Sigma t_{bj} Y_{b..j} + \Sigma\Sigma\Sigma t_{lhik} Y_{lhik} - \left\{ m'Y_{....} + \Sigma r_h' Y_{.h..} + \Sigma c_i' Y_{..i.} \right\} \qquad \text{(III-3)}$$

$$= 1235 - \left\{ 8.8918919(113) + [-1.4932432(34) + .0067568(40) + 1.4864865(39)] + [-2.5135135(27) - 1.8918919(21) + 4.4054054(65)] \right\}$$

$$= 1235 - 1191.0135180 = 43.9864820$$

Column (ignoring entry, eliminating row) ss (2 d.f.):

$$C = m'Y_{....} + \Sigma r_h' Y_{.h..} + \Sigma c_i' Y_{..i.} - \Sigma \frac{Y_{.h..}^2}{b+n_{1h.}} \qquad \text{(III-4)}$$

(where $n_{1h.} = \sum_i n_{1hi}$)

$$= 1191.0135180 - \left(\frac{34^2 + 40^2 + 39^2}{4}\right)$$

$$= 1191.0135180 - 1069.25 = 121.7635180.$$

As a partial check,

$$121.7635180 + 5.1666667 + 43.9864820 = 170.9166667$$

$$= \text{ss due to } m,\ r_h,\ c_i,\ t_{bj}, \text{ and } t_{1hik} - \frac{Y_{....}^2}{b^2 + \Sigma\Sigma n_{1hi}} \qquad \text{(III-5)}$$

$$= 1235 - \frac{113^2}{12}.$$

Row (ignoring entry, eliminating column effect) ss (2 d.f.):

$$m'Y_{....} + \Sigma r_h' Y_{.h..} + \Sigma c_i' Y_{..i.} - \Sigma \frac{Y_{..i.}^2}{b+n_{1.i}} \qquad \text{(III-6)}$$

$$= 1191.0135180 - \left\{ \frac{27^2}{4} + \frac{21^2}{3} + \frac{65^2}{5} \right\}$$

$$= 1191.0135180 - 1174.25 = 16.7635180$$

As a partial check,

$$16.7635180 = 5.1666667 - (110.1666667 - 121.7635180).$$

Also, the following sum of squares is of interest as a partial check on the other calculations:

Row + column (eliminating entry effect) ss (4 d.f.):

$$mY_{\cdot\cdot\cdot\cdot} + \Sigma r_h Y_{\cdot h\cdot\cdot} + \Sigma c_i Y_{\cdot\cdot i\cdot} + \Sigma t_{bj} Y_{b\cdot\cdot j} + \Sigma\Sigma\Sigma t_{lhik} Y_{lhik}$$

$$- \left\{ \frac{\Sigma Y_{b\cdot\cdot j}^{2}}{b} + \Sigma\Sigma\Sigma Y_{lhik}^{2} \right\} \qquad \text{(III-7)}$$

$$= 1235 - 582 - 569 = 84,$$

which should be the row + column sum of squares from the analysis of variance on variety yields only. For the example, 6 + 78 = 84 as it should.

The number of figures to the right of the decimal is larger than warranted by the accuracy of most data. The excessive number was carried because it was desired to show certain relations among the sums of squares free of rounding errors.

## IV. Augmented Incomplete Latin Square Design

The $v_b$ varieties may be arranged in a latin square design with one or more rows omitted or added to the $v_b$ x $v_b$ = b x b latin square. The Youden square [see 2] is a special case of an incomplete latin square design. As in the augmented latin square design there are $n_{lhi}$ seedlings included with one of the varieties in the h'th row and i'th column. The randomization procedure for the augmented incomplete latin square design follows that for the augmented latin square design. There are r rows and $v_b$ = b columns.

To illustrate, suppose that the $v_b$ = 4 varieties (A,B,C,D) are arranged in a Youden square design, that r = 3 rows, that b = 4 columns, and that $n_1$ = 17 seedlings (1,2,...,17). An arrangement such as the following might be obtained:

| Row | Column 1 | | | Column 2 | | | Column 3 | | Column 4 | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | 1 | 2 | C | 3 | 4 | B | 5 | 6 | D |
| 2 | 7 | D | 8 | 9 | B | 10 | A | 11 | C | 12 |
| 3 | 13 | 14 | C | 15 | 16 | A | 17 | D | B | |

The rows and columns are set up to control the maximum amount of variation possible in the experimental area.

The analysis for this design may be obtained by the methods given for designs II and III. Here again, the analysis is somewhat simplified if the $n_{lhi}$ are all equal, and the precision is usually greater if the $n_{lhi}$ are equal or nearly so.

The general form of the analysis of variance is:

| Source of Variation | Degrees of Freedom |
|---|---|
| Row (ignoring entry and column) | $r - 1$ |
| Column (ignoring entry; eliminating row) | $b - 1$ |
| Entry (eliminating row and column) | $b + v_1 - 1$ |
| Variety | $b - 1$ |
| Seedling and variety vs. seedling | $v_1$ |
| Error | $(b - 1)(r - 2)$ |

As before, the estimated experimental error may be obtained from an analysis of the varieties only.

V. Augmented Incomplete Block Design

The $v_b$ varieties may be arranged in any of the several lattice designs [see 2, Chapters XI, XII, and XIII] with $n_{1hi}$ seedlings in the i'th incomplete block of the h'th replicate. Thus, there will be k varieties plus $n_{1hi}$ seedlings in the i'th incomplete block of the h'th replicate. The k varieties are randomly allotted to the $n_{1hi} + k$ plots in the incomplete block and the $v_1 = \sum_{hi}\sum n_{1hi}$ seedlings are randomly allotted to the remaining $v_1$ plots after the varieties have been assigned. The procedure for allocating the $v_b$ varieties in a lattice design is the one described for lattice designs in texts on Experiment Design [e.g., see 2].

To illustrate, suppose that $v_b = 4$ varieties (A,B,C,D) are arranged in a balanced lattice design with $k = 2$ and $b = 3$ and suppose that $v_1 = 11$. Then, the following arrangement might have been obtained:

| | Replicate I | | | | Replicate II | | | | Replicate III | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | 1 | 3 | C | 5 | A | B | D | B | C | A | 11 |
| | B | 2 | D | 4 | C | 6 | 7 | 8 | 9 | 10 | D | |
| Block No. | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | |

The analysis of variance for an augmented one-restrictional lattice design of the above type with $v_b = k^2$ varieties and with k varieties in each incomplete block is:

| Source of Variation | Degrees of Freedom |
|---|---|
| Replicate (ignoring variety and seedling) | $b - 1$ |
| Block (eliminating variety and seedling) | $b(k - 1)$ |
| Variety and seedling (ignoring block) | $k^2 + v_1 - 1$ |
| Variety | $k^2 - 1$ |
| Seedling and variety vs. seedling | $v_1$ |
| Intrablock error | $k^2(b - 1) - bk + 1$ |

The sums of squares for blocks (eliminating varieties and seedlings), for varieties (ignoring blocks), and for intrablock error may be computed in the manner for the lattice design used and from the yields of varieties only.

The incomplete block design may be a simple, triple, rectangular, cubic, etc., lattice or one of the lattice square designs. Also, one of the other incomplete block designs could be used [see 2, Chapters XI to XIII].

OTHER AUGMENTED DESIGNS

There exists the possibility of using a magic latin square, a graeco-latin square, a tied latin square, split plot, etc., for the varieties in the experiment. Then, the seedlings could be included in much the same manner as for Designs I to V.

Also, designs could be set up with $v_b$ varieties occurring once in each of the blocks, $v_2$ varieties appearing in two of the b blocks, and $v_1$ varieties appearing in one of the b blocks. For certain values of $v_2$, the $v_2$ and $v_1$ varieties could be included in an ordinary chain block design with the $v_b$ varieties appearing in every block. Such a design could be called an augmented chain block design. Also, additional augmented chain block designs could be set up utilizing the chain blocks described by Mandel [4].

In addition to the above designs, the experimenter may wish to use all of varying amounts of experimental material and all meaningful methods of stratification. As an example, consider the following amounts of material allocated to the plots in the b blocks:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $v_b$ | entries | appear | once | in | each of | the | b | blocks | , |
| $v_b-1$ | " | " | " | " | b-1 | " | " | " | , |
| $v_b-2$ | " | " | " | " | b-2 | " | " | " | , |
| ⋮ | | | | | | | | | |
| $v_2$ | " | " | " | " | two | " | " | " | , and |
| $v_1$ | " | " | " | " | one | " | " | " | . |

Such a design, although a relatively simple one in this group, would be rather laborious to analyze. Unless there is good reason for designing such an experiment and available personnel for analysis, one of the simpler designs discussed thus far should be used. The above example is relatively simple in that only one-way stratification is used. If two-or more-way stratification of the experimented area were used, the analysis becomes much more complex and involved.

DISCUSSION

Although these designs were developed primarily for comparing seedlings and varieties developed by the sugar cane, pineapple, etc., geneticists, they have much wider usage. The entomologist and pathologist may wish to test a large number of new chemical compounds each year. The first test on such material is usually to decide if the new material shows promise, and anything that is obviously bad is discarded. Usually, only one replicate of such material is used. Also, the agronomist or physiologist may wish to compare a large number of new chemicals as weed sprays. The chemist in the laboratory may desire an estimate of experimental error for this material; he may use one of the chain block designs proposed by Youden et al. [4,7,8] or one of the designs presented herein.

Since the analysis for the more complex augmented designs has not been developed, the question may arise as to what to do with seedlings for which there is enough available material to plant two plots. The seedling could be included once in two different tests or as two seedlings in one test. If the latter, the mean of the two adjusted means would be used as the seedling mean. This procedure is not efficient in that no use is made of the extra degree of freedom which could be allocated to error. This lack of efficiency will only be serious if the number of degrees of freedom for error is relatively low, say less than 10 to 16, or if several seedlings are repeated twice in each experiment. As described in the section under Other Augmented Designs, an analysis could be developed for the above situation but has not been to date.

The question of systematic placement of the $v_b$ varieties arises from time to time. If the varieties are systematically spaced, say every third plot, if the analysis prescribed herein is used, the resulting mean differences and the experimental error variances are biased. Since the systematic placement of varieties invalidates known analyses and since stratification, e.g. Design III versus Design I, accomplishes more than systematic placement of varieties, there appears to be no argument for using a systematic arrangement of varieties or "check plots" in seedling tests.

SUMMARY

In response to the need for more efficient designs for comparing seedlings in the early stages of a breeding program, a new class of experimental designs, augmented designs, was developed. The analyses for the augmented randomized complete block and the augmented latin square designs are given and illustrated with numerical examples. Basically, the designs presented herein have one set of treatments (the varieties) repeated b times and a second set of treatments (the seedlings) appearing only once. A standard design (except for additional plots in the block) and analysis are used for the varieties while the augmented design analysis is used for the varieties and seedlings together.

The designs are also useful in the fields of entomology, pathology, chemistry, physiology, agronomy, and perhaps others for combining screening experiments on new material and preliminary testing experiments on promising material.

## LITERATURE CITED

1. Eisenhart, C. On the Statistical Analysis of Linked-Block Experiments (abstract). Biometrics 7:125, 1951.

2. Federer, W. T. Experimental Design - Theory and Application. Macmillan, N. Y., 1955.

3. Federer, W. T. A Method for Evaluating Genetic Progress in a Sugar Cane Breeding Program. Hawaiian Planters' Record 56: , 1955.

4. Mandel, J. Chain Block Designs with Two-Way Elimination of Heterogeneity. Biometrics 10:251-272, 1954.

5. Mangelsdorf, A. J. Sugar Cane Breeding in Hawaii - Part II, 1921 - 1952. Hawaiian Planters' Record 54:101-137, 1953.

6. Warner, J. N. The Evolution of a Philosophy on Sugar Cane Breeding in Hawaii. Hawaiian Sugar Planters' Record 54:139-162, 1953.

7. Youden, W. J. Linked Blocks: A New Class of Incomplete Block Designs (abstract). Biometrics 7:124, 1951.

8. Youden, W. J. and Conner, W. S. The Chain Block Design. Biometrics 9:127-140, 1953.

APPENDIX A. Derivation of Results for an Augmented Randomized Complete Block Design.

For the augmented randomized complete block design, the following linear model is used:

$$Y_{hijg} = \mu + \rho_j + \tau_{bi} + \tau_{1jg} + \epsilon_{bij} + \epsilon_{1jg}, \tag{A1}$$

where $i = 1,2,\ldots, v_b$ = number of varieties repeated b times; $j = 1,2,\ldots,$ $b$ = number of blocks; $g = 1,2,\ldots, n_{1j}$ = number of seedlings in j'th block and where a particular seedling does not appear in any other block; $h = 1$ or $b$ to denote the effects associated with seedlings and the effects associated with varieties; $\mu$ = mean effect; $\rho_j$ = effect of j'th block; $\tau_{bi}$ = effect of i'th variety; $\tau_{1jg}$ = effect of g'th seedling in the j'th block; and $\epsilon_{1jg}$ and $\epsilon_{bij}$ are random effects from a population with mean zero and variance $\sigma_\epsilon^2$.

The least squares estimates of $\mu$, $\rho_j$, $\tau_{bi}$, and $\tau_{1jg}$ are denoted as $m$, $r_j$, $t_{bi}$, and $t_{1jg}$, respectively; the estimates $m$, $r_j$, $t_{bi}$, and $t_{1jg}$ are those which satisfy the following normal equations:

Equation for the mean effect m:

$$\left(\sum_{j=1}^{b} n_{1j} + bv_b\right) m + b\sum_{i=1}^{v_b} t_{bi} + \sum_{j=1}^{b}\sum_{g=1}^{n_{1j}} t_{1gj}$$

$$+ \sum_j (v_b + n_{1j}) r_j = Y_{\ldots} = \text{total sum of N items} \tag{A2}$$

Equations for block effects $r_j$:

$$(v_b + n_{11})(m + r_1) + \sum_{i=1}^{v_b} t_{bi} + \sum_{g=1}^{n_{11}} t_{11g} = Y_{.1.} = \text{1st block total} \tag{A3}$$

$$\vdots$$

$$(v_b + n_{1j})(m + r_j) + \sum_{i=1}^{v_b} t_{bi} + \sum_{g=1}^{n_{1j}} t_{1jg} = Y_{.j.} = \text{j'th block total} \tag{A4}$$

$$\vdots$$

$$(v_b + n_{1b})(m + r_b) + \sum_{i=1}^{v_b} t_{bi} + \sum_{g=1}^{n_{1b}} t_{1bg} = Y_{.b.} = \text{b'th block total} \tag{A5}$$

Equations for variety effects $t_{bi}$:

$$b(m + t_{b1}) + \sum_{j=1}^{b} r_j = Y_{1.} = \text{1st variety total.} \quad (A6)$$

$$\vdots$$

$$b(m + t_{bi}) + \sum_{j=1}^{b} r_j = Y_{i.} = \text{i'th variety total.} \quad (A7)$$

$$\vdots$$

$$b(m + t_{bv_b}) + \sum_{j=1}^{b} r_j = Y_{v_b.} = v_b\text{'th variety total.} \quad (A8)$$

Equations for seedling effects $t_{1jg}$:

$$m + t_{111} + r_1 = Y_{111} = \text{1st seedling total in block 1.} \quad (A9)$$

$$\vdots$$

$$m + t_{11n_{11}} + r_1 = Y_{11n_{11}} = n_{11}\text{'th seedling total in block 1.} \quad (A10)$$

$$m + t_{121} + r_2 = Y_{121} = \text{1st seedling total in block 2.} \quad (A11)$$

$$\vdots$$

$$m + t_{12n_{12}} + r_2 = Y_{12n_{12}} = n_{12}\text{'th seedling total in block 2.} \quad (A12)$$

$$\vdots$$

$$m + t_{1jg} + r_j = Y_{1jg} = \text{g'th seedling total in j'th block.} \quad (A13)$$

$$\vdots$$

$$m + t_{1bn_{1b}} + r_b = Y_{1bn_{1b}} = n_{1v_1} \text{ seedling total in b'th block.} \quad (A14)$$

Assumptions

$$\sum_{j=1}^{b} r_j = 0. \quad (A15)$$

$$\sum_{i=1}^{v_b} t_{bi} + \sum_{j=1}^{b} \sum_{g=1}^{n_{1j}} t_{1jg} = 0. \quad (A16)[1]$$

---

[1] The restrictions $\Sigma t_{bi} = 0$ and $\Sigma\Sigma t_{1jg} = 0$ could be used instead of (A16) but these are considered to be unrealistic assumptions; if they were used different estimates for the mean, $m^*$, and of the blocks effects, $r_j^*$, would be obtained.

The solution of the above equations results in the following:

$$m = \frac{1}{v_b+v_1}\left\{ Y_{\cdot\cdot\cdot} - (b-1)\sum_{i=1}^{v_b}\frac{Y_{i\cdot}}{b} - \sum_{j=1}^{b}\frac{n_{1j}}{v_b}\left[ Y_{\cdot j\cdot} - \sum_{i=1}^{v_b}\bar{y}_{i\cdot} - \sum_{g=1}^{n_{1j}} Y_{1jg}\right]\right\}, \tag{A17}$$

which for $n_{1j}$ = a constant reduces to

$$m = \frac{Y_{\cdot\cdot\cdot} - (b-1)\sum_{i=1}^{v_b} Y_{i\cdot}/b}{v_b + v_1} . \tag{A18}$$

$$t_{bi} + m = Y_{i\cdot}/b = \bar{y}_{i\cdot} . \tag{A19}$$

$$r_j = \frac{1}{v_b}\left\{ Y_{\cdot j\cdot} - \sum_{i=1}^{v_b}\bar{y}_{i\cdot} - \sum_{g=1}^{n_{1j}} Y_{1jg}\right\} . \tag{A20}$$

$$t_{1jg} + m = Y_{1jg} - r_j \tag{A21}$$

For the comparison of the mean of the varieties with the mean of the seedlings the following two estimates are required:

$$\frac{1}{v_b}\sum_{i=1}^{v_b} t_{bi} = \frac{1}{bv_b}\sum_{i=1}^{v_b} Y_{i\cdot} - m \tag{A22}$$

and

$$\frac{1}{v_1}\sum_{j=1}^{b}\sum_{g=1}^{n_{1g}} t_{1jg} = \frac{1}{v_1}\sum_{j=1}^{b}\sum_{g=1}^{n_{1j}} Y_{1jg} - m - \frac{1}{v_1}\sum_{j=1}^{b} n_{1j} r_j, \tag{A23}$$

where

$$v_1 = n_{1\cdot} = \sum_{j=1}^{b} n_{1j} . \tag{A24}$$

However, it is impossible to obtain estimates of $r_j$ without having estimates of $t_{1jg}$ or at least $\sum_{g=1}^{n_{1j}} t_{1jg}$. Thus, this sum of squares does not appear to be obtainable as such. If $n_{1j}$ = a constant, then the ordinary method for comparing the mean of the varieties with the mean of the seedlings would be appropriate since this sum of squares would be unaffected by the block differences.

It should also be noted that the comparisons among seedlings within the same block may be made in the usual manner since this comparison is unaffected by block differences. The sum of squares for the $n_{1j}$ seedlings in block j is

$$\sum_{g=1}^{n_{1j}} Y_{1jg}^{2} - (\Sigma Y_{1jg})^2/n_{1j}, \text{ with } n_{1j} - 1 \text{ degrees of freedom.} \tag{A25}$$

The variance of a difference between two variety means, say 1 and 2, is

$$\begin{aligned} V(\bar{y}_{1\cdot} - \bar{y}_{2\cdot}) &= E[\bar{y}_{1\cdot} - \bar{y}_{2\cdot}]^2 - (E[\bar{y}_{1\cdot} - y_{2\cdot}])^2 \\ &= E[\frac{1}{b}\sum_{j=1}^{b}(\mu + \rho_j + \tau_{b1} + \epsilon_{b1j} - \mu - \rho_j - \tau_{b2} - \epsilon_{b2j})]^2 \\ &= (E[\frac{1}{b}\Sigma(\mu + \rho_j + \tau_{b1} + \epsilon_{b1j} - \mu - \rho_j - \tau_{b2} - \epsilon_{b2j}])^2 \\ &= \frac{1}{b^2} E[\sum_j(\tau_{b1} + \epsilon_{b1j} - \tau_{b2} - \epsilon_{b2j})]^2 - \frac{1}{b^2}([\Sigma(\tau_{b1} - \tau_{b2})])^2 \\ &= \frac{1}{b^2}\left\{b^2(\tau_{b1} - \tau_{b2})^2 + b\sigma_\epsilon^2 + b\sigma_\epsilon^2 - b^2(\tau_{b1} - \tau_{b2})^2\right\} \\ &= \frac{2\sigma_\epsilon^2}{b}, \end{aligned} \tag{A26}$$

where formula (A1) is used and where the following expectations are assumed :

$$E[\tau_{bi}]^2 = \tau_{bi}^2 \tag{A27}$$

$$E[\epsilon_{bij}]^2 = \sigma_\epsilon^2 \tag{A28}$$

$$E[\epsilon_{bij}\epsilon_{buv}] = 0 \text{ for } i,j \neq u,v \tag{A29}$$

$$E[\epsilon_{bij}] = 0 \tag{A30}$$

$$E[\tau_{bi}\epsilon_{bij}] = 0 \tag{A31}$$

The variance of a difference between two seedling means, $Y_{1jg}'$, adjusted for block effects depends on the two seedlings involved. If the two seedlings, say 1 and 2, and in the same block, the variance is

$$V(Y_{1j1} - Y_{1j2}) = V(Y_{1j1}' - Y_{1j2}')$$
$$= E[\mu + \rho_j + \tau_{1j1} + \epsilon_{1gj} - \mu - \rho_j - \tau_{1j2} - \epsilon_{1j2}]^2$$
$$- (E[\mu + \rho_j + \tau_{1j1} + \epsilon_{1j1} - \mu - \rho_j - \tau_{1j2} - \epsilon_{1j2}])^2$$
$$= 2\sigma_\epsilon^2, \qquad (A32)$$

where it is assumed that the expectations in (A27) to (A31) hold for the $\tau_{1jg}$ substituted for the $\tau_{bi}$ and for the $\epsilon_{1jg}$ substituted for the $\epsilon_{bij}$. This assumes that the seedlings and varieties have the same error variance. If this assumption does not hold, then it will be necessary to replicate the seedlings in order to estimate the error variance for seedlings.

The variance of a difference between two adjusted seedling means, say $Y_{112}' - Y_{121}'$, for two seedlings not appearing in the same block is:

$$V(Y_{112}' - Y_{121}') = V(Y_{112} - r_1 - Y_{121} + r_2)$$
$$= V(Y_{112} - Y_{121} - \frac{1}{v_b}\left\{Y_{.1.} - Y_{.2.}\right\} + \frac{1}{v_b}\left\{\sum_{g=1}^{n_{11}} Y_{11g} - \sum_{g=1}^{n_{12}} Y_{12g}\right\})$$
$$= E[\mu + \rho_1 + \tau_{112} + \epsilon_{112} - \mu - \rho_2 - \tau_{121} - \epsilon_{121}$$
$$- \frac{1}{v_b}\left\{(v_b + n_{11})(\mu + \rho_1) + \sum_i(\tau_{bi} + \epsilon_{bi1}) + \sum_{g=1}^{n_{11}}(\tau_{11g} + \epsilon_{11g})\right.$$
$$\left. - (v_b + n_{12})(\mu + \rho_2) - \sum_i(\tau_{bi} + \epsilon_{bi2}) - \sum_{g=1}^{n_{12}}(\tau_{12g} + \epsilon_{12g})\right\}$$
$$+ \frac{1}{v_b}\left\{\sum_{g=1}^{n_{11}}(\mu + \rho_1 + \tau_{11g} + \epsilon_{11g}) - \sum_{g=1}^{n_{12}}(\mu + \rho_2 + \tau_{12g} + \epsilon_{12g})\right\}]^2$$
$$- \left\{E[\tau_{112} - \tau_{121} + \epsilon_{112} - \epsilon_{121} - \frac{1}{v_b}\sum_i(\epsilon_{bi1} - \epsilon_{bi2})]\right\}^2$$
$$= 2(v_b + 1)\sigma_\epsilon^2/v_b. \qquad (A33)$$

The variance of a difference between a variety and an adjusted seedling mean is

$$V(\bar{y}_{i.} - Y_{1jg}') = V(\bar{y}_{i.} - Y_{1jg} + \frac{1}{v_b}\left\{Y_{.j.} - \Sigma\bar{y}_{i.} - \sum_{g=1}^{n_{1j}} Y_{1jg}\right\})$$
$$= (1 + \frac{1}{b} + \frac{1}{v_b} + \frac{1}{bv_b})\sigma_\epsilon^2. \qquad (A34)$$

## APPENDIX B. Derivation of Results for an Augmented Latin Square Design.

For the augmented latin square design, the following linear model is used:

$$Y_{ghijk} = \mu + \rho_h + \gamma_i + \tau_{bj} + \tau_{1hik} + \epsilon_{bhij} + \epsilon_{1hik}, \quad \text{(B1)}$$

where $g = 1$ or $b$; $h = 1, 2, \ldots, b$; $i = 1, 2, \ldots, b$; $j = 1, 2, \ldots, b$ except that a variety only appears once in a row and once in a column; $k = 1, 2, \ldots, n_{1hi}$ = number of seedlings in h'th row and i'th column and any given seedling occurs in only one plot; $\mu$ = mean effect; $\rho_h$ = effect of h'th row; $\gamma_i$ = effect of i'th column; $\tau_{bj}$ = effect of j'th variety; $\tau_{1hik}$ = effect of k'th seedling in the h'th row and i'th column; and $\epsilon_{bhij}$ and $\epsilon_{1hik}$ are random effects from a population with mean zero and variance $\sigma_\epsilon^2$.

The least squares estimates of $\mu$, $\rho_h$, $\gamma_i$, $\tau_{bj}$, and $\tau_{1hik}$ are denoted as $m$, $r_h$, $c_i$, $t_{bj}$, and $t_{1hik}$, respectively; these estimates satisfy the following normal equations:

Equation for mean effect, m

$$\left(\sum_{hi}\sum n_{1hi} + b^2\right) m + b \sum_{j=1}^{b} t_{bj} + \sum_{h=1}^{b}\sum_{i=1}^{b}\sum_{k=1}^{n_{1hi}} t_{1hik}$$

$$+ \sum_{h=1}^{b}\left(b + \sum_{i=1}^{b} n_{1hi}\right) r_h + \sum_{i=1}^{b}\left(b + \sum_{h=1}^{b} n_{1hi}\right) c_i = Y_{\cdot\cdot\cdot\cdot\cdot} \quad \text{(B2)}$$

= grand total of all items.

Equations for row effects, $r_h$

$$\left(b + \sum_i n_{1hi}\right)(m + r_h) + \sum_{j=1}^{b} t_{bj} + \sum_{i=1}^{b}\sum_{k=1}^{n_{1hi}} t_{1hik} + \sum_i n_{1hi} c_i$$

$$= \sum_{i=1}^{b}\left\{\sum_{j=1}^{b} Y_{bhij} + \sum_{k=1}^{n_{1hi}} Y_{1hik}\right\} = Y_{\cdot h\cdot\cdot} \quad \text{(B3)}$$

= total of h'th row.

Equations for column effects, $c_i$

$$\left(b + \sum_h n_{1hi}\right)(m + c_i) + \sum_{j=1}^{b} t_{bj} + \sum_{h=1}^{b}\sum_{k=1}^{n_{1hi}} t_{1hik} + \sum_h n_{1hi} r_h$$

$$= \sum_h \left\{ \sum_{j=1}^{b} Y_{bhij} + \sum_{k=1}^{n_{1hi}} Y_{1hik} \right\} = Y_{..i.} \quad (B4)$$

= i'th column total.

<u>Equations for variety effects, $\hat{t}_{bj}$</u>

$$b(m + t_{bj}) + \sum_{h=1}^{b} r_h + \sum_{i=1}^{b} c_i = Y_{b..j} \quad (B5)$$

= j'th variety total.

<u>Equations for seedling effects, $\hat{t}_{1hik}$</u>

$$m + r_h + c_i + t_{1hik} = Y_{1hik} \quad (B6)$$

= yield for k'th seedling in the h'th row and i'th column.

<u>Restrictions</u>

$$\sum_{h=1}^{b} r_h = 0, \quad (B7)$$

$$\sum_{i=1}^{b} c_i = 0, \text{ and} \quad (B8)$$

$$\sum_{i=1}^{b} t_{bj} + \sum_{h=1}^{b} \sum_{i=1}^{b} \sum_{k=1}^{n_{1hi}} t_{1hik} = 0. \quad (B9)$$

The solution of the above equations results in the following estimates of effects:

$$m = \frac{1}{v_1 + b} \left\{ Y_{.....} - (b-1-\frac{2v_1}{b}) \Sigma \bar{y}_{b..j} - \frac{1}{b} \sum_{ih}\sum n_{1hi} Y_{bh..} - \frac{1}{b} \sum_{hi}\sum n_{1hi} Y_{b.i.} \right\}, \quad (B10)$$

where $v_1 = \sum_{ih}\sum n_{1hi}$, $Y_{bh..}$ = total of varieties in h'th row, and $Y_{b.i.}$ = total of varieties in i'th column.

$$\hat{r}_h = \frac{1}{b} \left\{ Y_{bh..} - \sum_j \bar{y}_{b..j} \right\}. \quad (B11)$$

$$c_i = \frac{1}{b} \left\{ Y_{b.i.} - \sum_j \bar{y}_{b..j} \right\}. \quad (B12)$$

$$t_{bj} + m = Y_{b..j/b} = \bar{y}_{b..j} \; . \tag{B13}$$

$$t_{1hik} + m = Y_{1hik} - r_h - c_i . \tag{B14}$$

The variance of a difference between two variety means, say 1 and 2, is $V(\bar{y}_{b..1} - \bar{y}_{b..2}) = E[\frac{1}{b}(b\mu + b\tau_{b1} + \Sigma\rho_h$

$$+ \Sigma\gamma_i + \underset{hi}{\Sigma\Sigma}\epsilon_{bhi1} - b\mu - b\tau_{b2} - \Sigma\rho_h$$

$$- \Sigma\gamma_i - \underset{hi}{\Sigma\Sigma}\epsilon_{bhi2})]^2 - \left\{\tau_{b1} - \tau_{b2}\right\}^2 = 2\sigma_\epsilon^{\,2}/b, \tag{B15}$$

where the experimental error $E_e$ is an estimate of $\sigma_\epsilon^{\,2}$.

The variance of a difference between adjusted means of two seedlings which appear together in the h'th row and i'th column, say 1hi1 and 1hi2, is $V(Y_{1hi1}' - Y_{1hi2}') = 2\sigma_\epsilon^{\,2}$. (B16)

The variance of a difference between the adjusted means of two seedlings which appear in the same column but different rows (or in the same row but different columns) is:

$$V(Y_{11i1}' - Y_{12i2}') = 2(1 + \frac{1}{b})\sigma_\epsilon^{\,2}. \tag{B17}$$

The variance of a difference between adjusted means of two seedlings appearing in different rows and different columns, say 1111 and 1222, is $V(Y_{1111}' - Y_{1222}') = (2 + \frac{4}{b})\sigma_\epsilon^{\,2}$. (B18)

The variance of a difference between a variety mean and an adjusted seedling mean, say variety j and seedling 111, is $V(\bar{y}_{b..j} - Y_{1111}')$ $= (1 + \frac{3}{b} - \frac{2}{b^2})\sigma_\epsilon^{\,2}$. (B19)