BALANCED LATIN SQUARES

by

Gregory Campbell and Susan Geller
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeoseries #80-26

September, 1980

BALANCED LATIN SQUARES

by

Gregory Campbell
Department of Statistics, Purdue University

and

Susan C. Geller
Department of Mathematics, Purdue University

SUMMARY

A more general concept of Latin squares than complete Latin squares is introduced. These balanced Latin squares are proved to exist for all orders. The random selection of complete and balanced squares is considered. The question of the F-distribution approximation of the randomization distribution over the collection of complete or balanced squares is addressed for Latin square designs and extended to cross-over designs. A link is provided to complete sets of orthogonal balanced Latin squares where they exist.

Some Key Words: complete Latin squares, crossover designs, changeover designs, randomization distribution, orthogonal Latin squares.

Balanced Latin Squares

by

G. Campbell and S. Geller

1. INTRODUCTION: Historically, Latin square designs were developed to eliminate row and column fertility gradients in agriculture field experiments. However, fertility gradients need not be only horizontal or vertical. To guard against diagonal gradients too, Knut Vik designs were introduced. An $n \times n$ Latin square is Knut Vik if, when a replicate of the square is placed next to it, all main diagonals and off diagonals of length n of the $n \times 2n$ rectangle contain each treatment symbol once. For $n = 7$, there are exactly four Knut Vik designs, as mentioned in Atkin, Hay, and Larson (1977). Letting the integers 1 to 7 denote the treatments, they are:

1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
6 7 1 2 3 4 5	5 6 7 1 2 3 4	4 5 6 7 1 2 3	3 4 5 6 7 1 2
4 5 6 7 1 2 3	2 3 4 5 6 7 1	7 1 2 3 4 5 6	5 6 7 1 2 3 4
2 3 4 5 6 7 1	6 7 1 2 3 4 5	3 4 5 6 7 1 2	7 1 2 3 4 5 6
7 1 2 3 4 5 6	3 4 5 6 7 1 2	6 7 1 2 3 4 5	2 3 4 5 6 7 1
5 6 7 1 2 3 4	7 1 2 3 4 5 6	2 3 4 5 6 7 1	4 5 6 7 1 2 3
3 4 5 6 7 1 2	4 5 6 7 1 2 3	5 6 7 1 2 3 4	6 7 1 2 3 4 5

Note that each design is a row permutation of the others. One objection, most forcefully voiced by R. A. Fisher (1951), is to the exclusive use of systematic designs rather than randomly chosen designs. Proper randomization from all Latin squares of order $n \geq 5$ leads to a statistic whose randomization distribution conditional on the observations is well-approximated by the F-distribution of the normal model. R. A. Fisher was especially worried that restricting selection to a small number of designs invites a conniving nature to choose a pattern for the treatments which might cause the experimenter to falsely reject the null hypothesis of equality of treatments. Also, the distribution of the statistic from a small number of designs does not closely approximate the F-distribution. For further discussion see section three.

Another problem is that Knut Vik designs have an undesirable property; namely, each treatment is bordered by exactly the same pattern of other treatments. For example, in the first 7 × 7 square above, 1 is always bordered by the pattern $\begin{matrix} & 3 & \\ 7 & & 2 \\ & 6 & \end{matrix}$. In particular, in the rows of the first square 1 precedes 2 six times. In the event of some border interactions, this design could mask real effects.

Williams (1949) proposed the concept of completeness in order to balance the effects of neighboring treatments. A Latin square is row complete if for each element i, the element j, j ≠ i, immediately follows i exactly once in the n rows of the square. Column completeness is defined analogously. A Latin square is complete if it is both row and column complete. Note that the concept of a complete Latin square is different from a complete set of orthogonal Latin squares. Williams (1949) constructed a row complete Latin square for each even order. These squares can be rearranged by row permutations to form a complete Latin square. Complete squares exist for some odd orders; e.g., for n = 21 see Dénes and Keedwell (1974). But they do not exist for others; e.g., n = 3, 5, 7, see Hedayat and Afsarinejad (1978).

Dénes and Keedwell (1974) conjectured that every row complete Latin square can be transformed by row permutations to a complete square. Owens (1976) provided counterexamples for n = 10 and n = 14. The conjecture can also be disproved for odd order squares by using exhaustion on the 9 × 9 row complete square of K. B. Martz in Hedayat and Afsarinejad (1978).

In none of the papers on complete squares is there a discussion of randomization or an enumeration of all such squares for a particular order.

Since complete squares need not exist for a given order, a more general notion of equalization of pairwise boundaries is introduced here. This idea, called balanced Latin squares, is defined in section 2 and some applications are presented. It is noted that complete squares are balanced, but there are balanced squares which are not complete. The theorem that balanced squares exist for all odd orders ≥ 3 is proved constructively. The construction also suggests a relationship between even complete squares and odd balanced ones.

In section 3 criteria are given to assure that the randomization distribution of the test statistic for a large set of Latin squares approximates an F-distribution. It is shown that the criteria are met for the set of all balanced squares of order 3 and for the set of all balanced squares of order 5. The randomization procedure is presented. Some complete sets of orthogonal balanced Latin squares are given which also satisfy the criteria. Conjectures are made for higher-order balanced squares. Designs that are merely row complete or row balanced are also considered.

In the final section, the method suggested by Williams (1949) of using pairs of squares to achieve completeness for odd orders is pursued and the balanced squares are utilized to extend this work. Also a more general idea of balance is suggested. The paper concludes with a general discussion of Latin squares, their usefulness and their limitations.

2. BALANCED LATIN SQUARES: A Latin square is row balanced if, for any treatment i, treatment j, $j \neq i$, is adjacent to i exactly twice in the n rows. It is clear that this is weaker than row complete since, if i immediately precedes j for all $i \neq j$, then i and j are row adjacent exactly twice, once as ij and once as ji.

The following example illustrates the usefulness of balanced squares. There are five machines arranged in a row and five operators, one for each machine. In the five experimental runs, each operator operates each machine once. It may be important which operators are located next to each other for each run, although the side is not important. To test whether there is a difference among the machines and among the operators, assuming no machine-operator interaction, a Latin square design is used. The columns denote the machines, the treatments, numbered, 1 to 5, denote the operators, and the rows are the five runs. One such balanced design is

1	2	3	4	5
2	4	1	5	3
4	5	2	3	1
5	3	4	1	2
3	1	5	2	4

Note that, for example, 2 is adjacent to 1 exactly twice but always on the same side. This suggests the usefulness of row balanced Latin squares in cross-over or change-over designs in which each treatment occurs in each of n positions and any treatment pair are adjacent in exactly two rows, although the order of the pair is not crucial.

A Latin square is column balanced if, for any treatment i, treatment j, $j \neq i$, is adjacent to i in exactly two columns. For simplicity a Latin square is called balanced if it is both row and column balanced. One possible use for a balanced Latin square design would be in a controlled indoor agricultural experiment in which wind and light are controlled to be non-directional. There is, nonetheless, a possibility of an interaction between plots with a common boundary. Since the order of the plots does not affect the interaction, a balanced design would be appropriate. One such design of order 5 is

1	2	3	4	5
2	4	5	3	1
3	5	2	1	4
4	3	1	5	2
5	1	4	2	3

Note that this design is balanced, in contrast to the previous design which was only row balanced.

It is clear that complete Latin squares are balanced, and hence balanced squares exist for all even orders. Row permutations of an even complete square may destroy completeness but possibly preserve balance. For example with $n = 6$, square A is complete while B, which is a row permutation of A, is balanced.

1	2	3	4	5	6
2	4	1	6	3	5
3	1	5	2	6	4
4	6	2	5	1	3
5	3	6	1	4	2
6	5	4	3	2	1

A

1	2	3	4	5	6
2	4	1	6	3	5
5	3	6	1	4	2
4	6	2	5	1	3
6	5	4	3	2	1
3	1	5	2	6	4

B

This suggests that, for even orders, there are more balanced squares than complete ones.

One question of interest is the existence of balanced Latin squares of odd orders. For square $L = \{\ell_{ij}\}$ of order n call the ℓ_{ii} the main diagonal and the $\ell_{i,n+1-i}$ the off diagonal, as i ranges from 1 to n.

<u>Theorem 1</u>: Balanced Latin squares exist for all odd orders.

<u>Proof</u>: Let ℓ_{ij} denote the entries of an $n \times n$ square matrix where $n = 2m + 1$. The entries denote the treatments which are labelled 1 to n. For $i \leq j$ define ℓ_{ij} as follows:

1. If $i \equiv j \pmod 2$, then

$$\ell_{ij} = \begin{cases} i + j & \text{if } i + j < n + 1,\ i \equiv 0 \pmod 2 \\ i + j - 1 & \text{if } i + j < n + 1,\ i \equiv 1 \pmod 2 \\ n & \text{if } i + j = n + 1 \\ n - k & \text{if } i + j > n + 1,\ i \equiv 0 \pmod 2 \\ n - k + 1 & \text{if } i + j > n + 1,\ i \equiv 1 \pmod 2 \end{cases}$$

where $k = i + j - (n + 1)$.

2. If $i \not\equiv j \pmod 2$, then

$$\ell_{ij} = \begin{cases} j - i & \text{if } i \equiv 0 \pmod 2 \\ j - i + 1 & \text{if } i \equiv 1 \pmod 2). \end{cases}$$

Complete the matrix by symmetry about the main diagonal; i.e.,

$\ell_{ij} = \ell_{ji}$.

The construction can also be described as follows:

Step 1: Write 1 2 ... n as the first row; i.e., $\ell_{1j} = j$.

Step 2: For i even in the first row, fill in the diagonal from upper left to lower right starting with i and alternating with i - 1.

Step 3: For i odd and less than n in the first row fill in the diagonal from upper right to lower left beginning with i and alternating with i+1.

Step 4: Fill in n for the main off diagonal; i.e., $\ell_{i,n+1-i} = n$.

Step 5: For the last column write n n-2 n-1 n-3 ... 1 2.

Step 6: For the even entries of the last column, fill in the diagonal from upper left to lower right beginning with i and alternating with i - 1.

Step 7: Complete by symmetry about the main diagonal.

For n = 5, the following is obtained:

1	2	3	4	5
2	4	1	5	3
3	1	5	2	4
4	5	2	3	1
5	3	4	1	2

It remains to prove that the square just defined is Latin and balanced.

To see that the square is Latin, write the i^{th} row. For $i \leq m = (n-1)/2$, i even, it is $i, i + 2, i - 2, i + 4, i - 4, \ldots, 2i - 2, 2, 2i, 1, 2i + 2, 3, 2i + 4, \ldots, 2m, 2m + 1 - 2i, 2m + 1, 2m + 3 - 2i, 2m - 1, \ldots, 2m - i - 1, 2m - i + 3, 2m - i + 1$. It is clear that every integer from 1 to n occurs exactly once. There are three other cases; namely, $i \leq m$, i odd; $i \geq m$, i even; and $i \geq m$, i odd. For each case, the rows can also be enumerated, noticing the occurrence of each treatment exactly once in each row. By symmetry about the main diagonal, each treatment occurs precisely once in each column. Thus the square is Latin.

The proof that it is balanced proceeds by induction. Note that, due to symmetry, it suffices to show that for entries ℓ_{ij}, $i \leq j$, each treatment borders every other treatment precisely once in the rows and once in the columns. An outline of the proof follows. The theorem is easily checked for n = 3 and 5. Suppose the 2m - 1 by 2m - 1 square is balanced. We shall construct the 2m + 1 by 2m + 1 square from the 2m - 1 by 2m - 1 square. Consider the portion of the 2m - 1 by 2m - 1 square ℓ_{ij}, $i \leq j$,

$i + j \leq 2m + 1$. For $2m - 1 = 9$ this is

1	2	3	4	5	6	7	8	9
	4	1	6	3	8	5	9	7 .
		5	2	7	4	9	6	
			8	1	9	3		
				9	2			

Replace every other $2m - 1$ by $2m$ beginning in row 2. Now take the portion of the square ℓ_{ij}, $i \leq j$, $i + j \geq 2m - 1$ and replace alternate $2m - 1$ by $2m$ starting in row 1 as illustrated for $2m - 1 = 9$:

			8	10
		5	9	7
	4	10	6	8
1	9	3	7	5
10	2	8	4	6
	7	1	5	3
		6	2	4
			3	2
				1

Place these two triangles together with an off diagonal of $2m + 1$'s between them, filling in the needed $2m - 1$ in the last column and $2m$ in the last row; e.g.

1	2	3	4	5	6	7	8	9	10	11
	4	1	6	3	8	5	10	7	11	9
		5	2	7	4	9	6	11	8	10
			8	1	10	3	11	5	9	7
				9	2	11	4	10	6	8
					11	1	9	3	7	5
						10	2	8	4	6
							7	1	5	3
								6	2	4
									3	1
										2

Note that the outlined pieces are just the part of the balanced 2m - 1 by 2m - 1 square in which all treatment pairs occur precisely once vertically and once horizontally. The pairs involving 2m + 1 with 2m and 2m - 1 are in the upper right corner. It is easy to prove that the elements protruding from the top toward the diagonal are 2, 3, 6, 7, 10, 11,... while those protruding from the bottom are 1, 4, 5, 8, 9,... . These are systematically bordered by 2m and 2m + 1 precisely once horizontally and once vertically. Thus the square is balanced. □

It is possible to construct complete even squares from odd balanced ones and vice versa. To go from the balanced square of order 2m + 1 of Theorem 1 to a complete square of order 2m, merely take the triangle of entries on and above the main diagonal and above the off diagonal; i.e., ℓ_{ij} such that $i \leq j$ and $i + j < n + 1$. Then symmetrize about both diagonals. For example, for 2m + 1 = 5,

$$\begin{matrix} 1 & 2 & 3 & 4 \\ & 4 & 1 & \\ & & & \\ & & & \end{matrix} \quad \rightarrow \quad \begin{matrix} 1 & 2 & 3 & 4 \\ & 4 & 1 & 3 \\ & & 4 & 2 \\ & & & 1 \end{matrix} \quad \rightarrow \quad \begin{matrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \\ 3 & 1 & 4 & 2 \\ 4 & 3 & 2 & 1 \end{matrix}$$

It can be easily shown that the resulting even complete square is the standard one of Dénes and Keedwell (p. 82, 1974) and Bradley (1958). Thus this gives an easier formula for the 2m by 2m complete square.

The other direction requires more effort. Write down the portion of the 2m by 2m complete design on and above both the main and off diagonals; i.e. ℓ_{ij} such that $i \leq j$ and $i + j \leq 2m+1$. Consider secondly the portion symmetric to it about the off diagonal and including the off diagonal; i.e., ℓ_{ij} such that $i \leq j$ and $i + j \geq 2m + 1$. In this second portion interchange 1 and 2, 3 and 4, etc. Attach this second portion to the first with an off diagonal of 2m + 1 and complete by symmetry about the main diagonal. For example, for 2m = 4 we get

$$
\begin{array}{cccc} 1 & 2 & 3 & 4 \\ & 4 & 1 & \end{array}
\;+\;
\begin{array}{cc} & 3 \\ 2 & 4 \\ 3 & 1 \\ & 2 \end{array}
\;\rightarrow\;
\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ & 4 & 1 & 5 & 3 \\ & & 5 & 2 & 4 \\ & & & 3 & 1 \\ & & & & 2 \end{array}
\;\rightarrow\;
\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 5 & 3 \\ 3 & 1 & 5 & 2 & 4 \\ 4 & 5 & 2 & 3 & 1 \\ 5 & 3 & 4 & 1 & 2 \end{array} .
$$

3. <u>RANDOMIZATION OF COMPLETE AND BALANCED SQUARES</u>: An important feature in ordinary Latin squares is the ability to select one at random from the collection of all Latin squares of a particular order. Conditional on any given set of observations, it is possible to calculate the exact randomization distribution of the test statistic under the null hypothesis of no treatment effects by evaluation of the statistic for each of the equally likely Latin squares in the randomization. The advantage of randomization is that, for $n \geq 5$, this randomization distribution is well-approximated by the F-distribution. Under the assumption of normally distributed errors, the test statistic has exactly this F-distribution. For this reason Fisher (1951) has been extremely vociferous in promoting complete randomization in the Latin square designs. This entails random selection of a standard square; i.e. one with 1 2 ... n as first column and first row. Then randomization of the rows of this square excluding the first, and lastly randomization of the treatments completes the selection. In the case of order 6 or larger, one standard square is given and the randomization is on all rows, columns and treatments. A natural question then is whether randomization is viable for complete or balanced squares and, if so, whether the randomization distribution of the test statistic is closely approximated by the F-distribution.

First note that row balance and row completeness are preserved by row permutations, but column permutations can destroy these row properties. Thus row or column permutations can destroy balance or completeness. One can nonetheless permute treatments without affecting any of these properties. Unfortunately the value of the test statistic is unchanged by treatment permutations. Thus randomization requires enumeration of balanced or complete squares, excluding treatment permutation.

First, however, what properties do a large set of squares require to insure adequate approximation of the randomization distribution by an F-distribution? As in Scheffé (1959), design random variables are needed. Let $d_{ijk} = 1$ if treatment k occurs in row i, column j, and 0 otherwise. Only the first and second moment conditions of Scheffé (1959) are required; namely,

1) $E(d_{ijk}) = 1/n$ and

$$2)\ E(d_{ijk}\, d_{i'j'k'}) = \begin{cases} 1/n & \text{if } i = i',\ j = j',\ k = k' \\ 0 & \text{if only 2 of } i = i',\ j = j',\ k = k' \text{ hold} \\ 1/n(n-1) & \text{if only 1 of } i = i',\ j = j',\ k = k' \text{ holds} \\ (n-2)/n(n-1) & \text{if } i \neq i',\ j \neq j',\ k \neq k'. \end{cases}$$

where the expectation is taken over the equally likely selection of a square from the set. It is immediate that any set of Latin squares which allows randomization on treatments is guaranteed to have $E(d_{ijk}) = 1/n$, and hence $E(d^2_{ijk}) = 1/n$ also. Furthermore, if only two of $i = i'$, $j = j'$, $k = k'$ hold, then $E(d_{ijk}\, d_{i'j'k'}) = 0$ since the squares are Latin. Thus the only constraints are those for which equality of the indices occurs exactly once or not at all. These can be reformulated as follows, as suggested by Cox (1980). Let $N_{(ij)(i'j')}(k,k')$ denote the number of squares for which treatment k is in position ij and k' in i'j'. Then the equivalent conditions are

a) $N_{(ij)(i'j')}(k,k) = c_1$ for fixed $i \neq i'$, $j \neq j'$ and for all k

b) $N_{(ij)(i'j')}(k,k') = c_2$ for fixed $i \neq i'$, $j \neq j'$ and all $k \neq k$.

Furthermore c_1 and c_2 do not depend on ij and i'j'. Moreover, if conditions 1) and 2) hold, since for every square with the same treatment in the two locations, $i \neq i'$, $j \neq j'$, there must be n - 2 other squares with n - 2 different pairs in those locations, $c_2 = (n-2)c_1$.

The natural way to proceed is to generate groups of n - 1 squares with a common first row such that the conditions are satisfied for each group. Complete sets of mutually orthogonal Latin squares are such groups.

Theorem 2: A complete set of n - 1 mutually orthogonal Latin squares of order n with common first row satisfies conditions a) and b).

Proof: Fix a pair of locations ij and i'j' such that $i < i'$ and $j \neq j'$. Label the squares 1, 2,...,n - 1 and let k_r and k_r' be the treatment in the r^{th} square in location ij and i'j' respectively. Recall that square s is orthogonal to square r if each treatment pair $k_r k_s$ occurs exactly once among all ij locations. First note that $k_r = k_r'$ for at most one r since, if $k_r = k_r'$ and $k_s = k_s'$ for $r \neq s$, then the pair $k_r k_s$ occurs twice in the enumeration of pairs from squares r and s, contradicting orthogonality. It is claimed that for each location pair, there is a square r such that $k_r = k_r'$. It suffices to show that for any Latin square 1/(n-1) of the pairs of locations with $i < i'$ have the same treatment, since there are n - 1 squares in the set and for a given pair of locations $k_r = k_r'$ at most once. In all, there are $n^2(n-1)^2/2$ pairs of locations, $i < i'$, $j \neq j'$, for a given square r. Holding i and i' constant, there are n pairs of locations such that $k_r = k'_r$. Hence for square r, there are $n \binom{n}{2} = n^2(n-1)/2$ such pairs. Thus for a given pair of locations, there is one square such that $k_r = k'_r$ and n - 2 squares such that $k_s \neq k'_s$. By orthogonality and $i < i'$, $k'_r \neq k'_s$ for $r \neq s$, and $k_r = k_s$ if and only if $i = 1$. Hence all treatment pairs (k,k') with $k \neq k'$ occur with the same frequency. Therefore conditions a) and b) hold with $c_2 = (n-2)c_1$. □

For n = 3, there are only 2 balanced squares before treatment randomization. They are row permutations of each other and orthogonal; thus they satisfy the above condition. In fact they are the only 3 by 3 Latin squares with 1 2 3 as the first row. They are

$$\begin{matrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{matrix} \quad \text{and} \quad \begin{matrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{matrix}.$$

However, this results in only two different values of the test statistic, far too few to allow adequate approximation by the F-distribution.

For n = 5 there are three standard balanced squares. Each such square generates by row randomization three complete sets of four orthogonal balanced squares. Thus there are nine sets of four squares which separately satisfy the conditions. See Table 1 for these squares, where each row of the table is a complete set of orthogonal balanced Latin squares. Note that these are all the balanced squares of order 5 up to treatment randomization. These give 36 different values of the test statistic which may be approximated crudely by the F-distribution. A preferred approach is to generate the exact null distribution under randomization of the test statistic conditional on the observed values in order to perform the test. The randomization procedure is to randomly select one of these 36 squares and then to randomly permute the treatment.

TABLE 1

1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 4 1 5 3	3 1 5 2 4	4 5 2 3 1	5 3 4 1 2
3 1 5 2 4	2 4 1 5 3	5 3 4 1 2	4 5 2 3 1
4 5 2 3 1	5 3 4 1 2	3 1 5 2 4	2 4 1 5 3
5 3 4 1 2	4 5 2 3 1	2 4 1 5 3	3 1 5 2 4
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 4 1 5 3	3 1 5 2 4	4 5 2 3 1	5 3 4 1 2
5 3 4 1 2	4 5 2 3 1	2 4 1 5 3	3 1 5 2 4
4 5 2 3 1	5 3 4 1 2	3 1 5 2 4	2 4 1 5 3
3 1 5 2 4	2 4 1 5 3	5 3 4 1 2	4 5 2 3 1
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 4 1 5 3	3 1 5 2 4	4 5 2 3 1	5 3 4 1 2
5 3 4 1 2	4 5 2 3 1	2 4 1 5 3	3 1 5 2 4
3 1 5 2 4	2 4 1 5 3	5 3 4 1 2	4 5 2 3 1
4 5 2 3 1	5 3 4 1 2	3 1 5 2 4	2 4 1 5 3
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 4 5 3 1	3 5 2 1 4	4 3 1 5 2	5 1 4 2 3
3 5 2 1 4	5 1 4 2 3	2 4 5 3 1	4 3 1 5 2
4 3 1 5 2	2 4 5 3 1	5 1 4 2 3	3 5 2 1 4
5 1 4 2 3	4 3 1 5 2	3 5 2 1 4	2 4 5 3 1
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 4 5 3 1	3 5 2 1 4	4 3 1 5 2	5 1 4 2 3
3 5 2 1 4	5 1 4 2 3	2 4 5 3 1	4 3 1 5 2
5 1 4 2 3	4 3 1 5 2	3 5 2 1 4	2 4 5 3 1
4 3 1 5 2	2 4 5 3 1	5 1 4 2 3	3 5 2 1 4
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 4 5 3 1	3 5 2 1 4	4 3 1 5 2	5 1 4 2 3
5 1 4 2 3	4 3 1 5 2	3 5 2 1 4	2 4 5 3 1
4 3 1 5 2	2 4 5 3 1	5 1 4 2 3	3 5 2 1 4
3 5 2 1 4	5 1 4 2 3	2 4 5 3 1	4 3 1 5 2
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 5 4 1 3	3 4 2 5 1	4 1 5 3 2	5 3 1 2 4
3 4 2 5 1	4 1 5 3 2	5 3 1 2 4	2 5 4 1 3
4 1 5 3 2	5 3 1 2 4	2 5 4 1 3	3 4 2 5 1
5 3 1 2 4	2 5 4 1 3	3 4 2 5 1	4 1 5 3 2
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 5 4 1 3	3 4 2 5 1	4 1 5 3 2	5 3 1 2 4
3 4 2 5 1	4 1 5 3 2	5 3 1 2 4	2 5 4 1 3
5 3 1 2 4	2 5 4 1 3	3 4 2 5 1	4 1 5 3 2
4 1 5 3 2	5 3 1 2 4	2 5 4 1 3	3 4 2 5 1
1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
2 5 4 1 3	3 4 2 5 1	4 1 5 3 2	5 3 1 2 4
4 1 5 3 2	5 3 1 2 4	2 5 4 1 3	3 4 2 5 1
5 3 1 2 4	2 5 4 1 3	3 4 2 5 1	4 1 5 3 2
3 4 2 5 1	4 1 5 3 2	5 3 1 2 4	2 5 4 1 3

For higher order squares the list of all standard squares is too long to check individually for balance or completeness. In addition, it is also extremely tedious to check all row permutations, keeping the first constant, for balance or completeness. Thus it would be preferable to have another method of generating sets of balanced or complete squares which satsify the above criteria. One such method is to concentrate on complete sets of orthogonal balanced or complete squares. It is well-known that complete sets of orthogonal squares exist for all primes and all prime powers. Furthermore some sets of each order can be generated by row permutation; see p. 167 and Chapter 7 of Dénes and Keedwell (1974). The randomization would then be from a small group of complete sets of orthogonal Latin squares.

Consider balanced squares first. It is as yet unresolved whether there are complete sets of orthogonal balanced squares for all powers of odd primes. However the following are some partial results. The general idea is to show that the standard balanced square of Theorem 1 is in the transformation set of the square G which is generated by the Galois field of order p (GF(p)) in such a way as to give a complete orthogonal set by row permutation; see p. 167 of Dénes and Keedwell (1974). To do this we need

<u>LEMMA 3</u>: For n odd, under the treatment permutation which interchanges treatments i and n - i + 1 for i odd and leaves even i fixed, the standard balanced square L in the proof of Theorem 1 is transformed to the square $A = \{a_{ij}\}$ such that $a_{ij} - a_{i1} \equiv a_{1j} \pmod{n}$ for all i and j.

Proof: The first row of the standard square L is transformed to 0, 2, n - 2, 4, n - 4, -3, n - 1, 1. Add, modulo n, the transformed value a_{i1} of $\ell_{i1} = i$ to each entry. The proof consists of showing that the inverse transformation of this row is then the i^{th} row of the original matrix L of Theorem 1. For example, if i is even and $\leq(n-1)/2$, the addition of the above row by i modulo n results in the row:

$$i, i + 2, i - 2, i + 4, i - 4, \ldots, i + 3, i - 1, i + 1$$

which when transformed back gives the row:

$$i, i + 2, i - 2, i + 4, i - 4, \ldots, n - i - 2, n - i + 2, n - i.$$

This is just the i^{th} row of L in Theorem 1. The remainder of the proof is to check the rows for the other three cases of Theorem 1.

Theorem 4: For odd primes p the standard balanced square of order p is in the transformation set of the square G generated by GF(p) (p. 167 Dénes and Keedwell (1974)).

Corollary 5: The standard balanced square B of order p, an odd prime, is a member of a complete orthogonal set.

Proof: By Theorem 4 there is a transformation taking G to B. Since orthogonality is preserved by transformations, the corollary follows. □

Proof of Theorem: By Dénes and Keedwell (1974) p. 167, the square G has the property $a_{ij} \equiv a_{i1} + a_{ij}$ (mod n) and hence has the property of the treatment permuted square in Lemma 3. GF(p) has elements 0, 1,...,p - 1. If the columns of G are permuted to the order 0 2 p - 2 ... 3 p - 1 1 and then the rows to the same order, the property is preserved. The result is the permuted square of Lemma 3. □

Note that the transformation does not work for p^s, $s \geq 2$ since the additive group of GF(p^s) is not cyclic. It is conjectured that the standard balanced square of order p^s, $s \geq 2$, p any prime, is not in the transformation set of G. Whether the balanced square is part of a complete orthogonal set for p^s odd is a good question.

It is also conjectured that, for p an odd prime, the squares of the complete orthogonal set of Corollary 5 are also balanced. This is true for p = 3, 4, 7, 11 and 13. Note that, since the orthogonal set of the Galois square is obtained by row permutation, so is the orthogonal set of the balanced square. Thus the other members of the complete orthogonal set are determined by the first column. For p = 7 and p = 11 the (p-1) orthogonal squares have first column as follows:

1	1	1	1	1	1
2	3	4	5	6	7
3	2	5	4	7	6
4	5	7	6	3	2
5	4	6	7	2	3
6	7	3	2	4	5
7	6	2	3	5	4

1	1	1	1	1	1	1	1	1	1
2	3	4	5	6	7	8	9	10	11
3	2	5	4	7	6	9	8	11	10
4	5	8	9	11	10	7	6	3	2
5	4	9	8	10	11	6	7	2	3
6	7	11	10	5	4	2	3	8	9
7	6	10	11	4	5	3	2	9	8
8	9	7	6	2	3	10	11	5	4
9	8	6	7	3	2	11	10	4	5
10	11	3	2	8	9	5	4	6	7
11	10	2	3	9	8	4	5	7	6

It is further conjectured that there are at least (p-2) complete sets of orthogonal balanced Latin square. This would nicely take care of odd order Latin squares up to and including 19 excluding 9 and 15. A further conjecture is that for p a prime that any balanced Latin square, although not necessarily generated by the regular standard one, is also in a complete set of balanced orthogonal Latin squares.

For even orders the case seems less encouraging. For n = 4 there are two squares up to treatment randomization which are complete and none which are balanced but not complete. Thus it is impossible to satisfy condition a). For n = 6 there are 19 balanced squares obtained by row permutations from the usual complete one. It is unknown whether there are more standard squares which are balanced or complete. Unfortunately no subset of the known 19 balanced squares satisfies condition a). Higher even order squares have not been investigated.

Consider the Latin square designs that require completeness or balance only in the rows or columns. For simplicity, restrict the discussion to properties of rows rather than columns. Such designs are particularly useful in cross-over designs, also called change-over designs. Note that row completeness and row balance are preserved under row permutation. Further, the value of the appropriate test statistic for such designs changes under row permutation but not under treatment permutation. In addition, the first and second moment conditions of Scheffé (1959) hold for sets of row-complete or row-balanced designs which include row permutation and treatment permutation due to the fact that the squares are Latin. Therefore, the randomization proceeds as follows for balanced squares, with a similar procedure for completeness. Choose at random one standard balanced Latin square from some set of such squares. Randomly permute the rows excluding the first and randomly permute the treatments. For each standard balanced square there are (n-1)! values of the test statistic, due to the row permutation. If k denotes the number of standard balanced squares in the collection, there are k(n-1)! possible values of the test statistic. If this number is large, the F-distribution will closely approximate the randomization distribution for any particular set of observations.

For example, for $n = 4$ there is only one of the four standard Latin squares that is row complete and hence row balanced; it is also complete. There are hence only $3! = 6$ possible values of the test statistic. Then, the randomization distribution can be exactly calculated. For $n = 5$, three of the 56 standard squares are row balanced; they are also balanced. Thus, there are $3 \cdot 4! = 72$ possible values of the test statistic under randomization. This can be approximated somewhat crudely by the F-distribution. For larger n it is not necessary to enumerate all standard row balanced or row complete squares. Thus, for $n = 6$, using only one standard row complete square there are $5! = 120$ possible values.

In cross-over designs, the number of rows need not be n; it can be any multiple of n. Row completeness and row balance can be defined for analogously for $kn \times n$ rectangles. The n! rows of all possible permutation of the n treatments is obviously row balanced and row complete as a $n! \times n$ rectangle. For $kn < n!$, row complete or row balanced designs can be constructed by randomly choosing k row complete or row balanced squares from the collection and then randomizing on rows and treatments.

4. CONCLUDING REMARKS: Unable to find a complete square of odd order, Williams (1940) constructed a pair of side by side squares such that row completeness is achieved as a pair. Houston (1966) called such squares (row) complimentary. It is possible to generate many such pairs of squares using balanced squares. For any balanced square of odd order, write the 180° rotation of that square to its right. Then the pair are row complimentary and column balanced. For example for $n = 5$

1 2 3 4 5	3 2 4 1 5
2 4 5 3 1	2 5 1 3 4
3 5 2 1 4	4 1 2 5 3
4 3 1 5 2	1 3 5 4 2
5 1 4 2 3	5 4 3 2 1

These squares as a pair are row and column complete.

The lack of complete success in section 3 with regard to even order Latin squares leads to more general notions of balance. One such concept is that of <u>overall balance</u>; i.e. for any treatment i, treatment j, $j \neq i$, borders it exactly four times in all on the horizontal and the vertical. There are squares which are overall balanced but not balanced. For $n = 4$ there are 5 such squares:

1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4
4 3 1 2	3 4 2 1	3 4 1 2	3 4 1 2	3 4 1 2
2 1 4 3	2 1 4 3	2 3 4 1	2 1 4 3	4 1 2 3
3 4 2 1	4 3 1 2	4 1 2 3	4 3 2 1	2 3 4 1

This type of square would be useful in such applications as indoor crop experiments with no directional preferences but possible boundary interactions in which there is no difference between horizontal and vertical boundaries. Unfortunately, for $n = 4$ no subset of overall balanced and complete squares satisfies the randomization criteria.

Latin squares have been a part of statistics at least since the nineteenth century. In many instances the economy of such designs in dramatic, requiring n^2 observations rather than the n^3 for a complete factorial. This paper has suggested the logical restriction to balanced designs as a way to avoid the possibility of a systematically faulty design. Such an approach is certainly superior to the practice of disguarding "bad" squares produced by randomization. One must keep in mind the limitations of Latin squares. If general interactions are known or suspected to exist, the Latin square design is inappropriate and either a fractional or complete factorial, either of which requires more observations, is recommended.

REFERENCES

Atkin, A.O.L., Hay, L., and Larson, R.G. (1977). Construction of Knut Vik designs. J. Statist. Plan. and Inf. 1, 289-297.

Box, J.F. (1978). R.A. Fisher, the Life of a Scientist. New York: John Wiley and Sons, Inc.

Bradley, J.V. (1958). Complete counterbalancing of immediate sequential effects in a Latin square design. J. of Amer. Statist. Assoc. 53 525-528.

Bugelski, B.R. (1949). A note on Grant's discussion of the Latin square principle in the design of experiments. Psych. Bull. 46, 49-50.

Cox, D.R. (1980). Personal communication.

Dénes, J. and Keedwell, A.D. (1974). Latin Squares and their Applications. New York: Academic Press.

Fisher, R.A. (1951). The Design of Experiments. Edinburgh: Oliver and Boyd Ltd. (seventh edition).

Fisher, R.A. and Yates, F. (1963). Statistical tables for biological, agricultural and medical research. Edinburgh: Oliver and Boyd Ltd. (sixth edition).

Hedayat, A. and Afsarinejad, K. (1978). Repeated measurement designs, II. Ann. Statist. 6 619-628.

Hedayat, A. and Federer, W. (1975). On the nonexistence of Knut Vik designs for all even orders. Ann. Statist. 3 445-447.

Houston, T.R. (1966). Sequential counterbalancing in Latin squares. Ann. Math. Statist. 37 741-743.

Owens, P.J. (1976). Solutions of two problems of Dénes and Keedwell on row-complete Latin squares. J. Comb. Th. (A) 21 299-308.

Scheffé, H. (1959). The Analysis of Variance. New York: Wiley.

Sharma, V.K. (1975). An easy method of constructing Latin square designs balanced for the immediate residual and other order effects. Can. J. Statist. 3 119-124.

Williams, E.J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. Australian J. of Sci. Research Ser. A, 2 149-168.

Williams, E.J. (1950). Experimental designs balanced for pairs of residual effects. Australian J. of Sci. Research Ser. A 3 351-363.