Quasi-Latin designs and their use in glasshouse experiments

C. J. Brien
University of South Australia, Adelaide, Australia.
R. A. Bailey
Queen Mary University of London, London, United Kingdom.
T. T. Tran
University of Adelaide, Adelaide, Australia.
J. Boland
University of South Australia, Adelaide, Australia.

[Received January 26, 2012]

Summary. This paper gives a general method for constructing quasi-Latin square, quasi-Latin rectangle and extended quasi-Latin rectangle designs for symmetric factorial experiments. Two further methods are given for parameter values satisfying certain conditions. Designs are constructed for a range of numbers of rows and columns so that the different construction techniques are illustrated. For some row and column combinations, different designs are compared, including designs constructed using computer search algorithms. The construction of designs with rows and columns that are nested or contiguous is also discussed.

Keywords: Factorial designs; Glasshouse experiments; Greenhouse experiments; Latinized designs; Quasi-Latin designs; Row-column designs

1. Introduction

The motivation for the work reported here comes from the need for designs for glasshouse experiments that involved several treatment factors (Tran, 2009), but is also applicable to experiments in other plant houses, such as in greenhouses, shade houses and polytunnels. In one experiment described by Tran (2009), a design was required for an experiment to investigate the effects of five treatment factors on the growth of species of Australian native plants that potentially could be used in the remediation of sites in the rail corridor either

E-mail: chris.brien@unisa.edu.au
Address for correspondence: Dr C.J. Brien, School of Mathematics and Statistics, University of South Australia, GPO Box 2741, Adelaide, SA 5001, Australia.

side of railway tracks. Three factors, each at two levels, were whether or not a nurse crop was used, whether or not gypsum was added to the soil and whether or not mycorrhizal fungi were added to the soil. The resulting eight treatments were to be applied to main plots that were arranged in a 4×10 rectangle. It was thought that there would be interactions between the factors and so it was important that the design gave good estimates of all interactions. The other two factors were to be applied to subplots and are not considered in this paper. A design for the main plot treatments is required and, as plots are arranged in a rectangle, designs such as the quasi-Latin square designs cannot be employed. Originally, Design 2 in Section 7.3 was obtained using CycDesigN 2.0 (Whitaker et al., 2002) to be used for the experiment, but the same design is produced using CycDesigN 4.0 (CycSoftware Ltd, 2009) with default settings. The confounding in this design was derived by inspection. Design 3 in Section 7.3 was then constructed by choosing suitable sets of confounding characters, after which a version of the algorithm in Section 4 was used. Before the experiment was run, the researchers decided to reduce the number of replicates of the 8 treatments from 5 to 3 and so Design 3 in Section 7.2 for a 4×6 rectangle was produced. It was employed in the experiment. The other designs in Sections 7.2 and 7.3 were constructed subsequently.

Usually glasshouses are carefully aligned on North/South and East/West axes as in Edmondson (1989) and Williams and John (1996) and this is the case for the glasshouse experiments described in Tran (2009). It is usually anticipated that, in glasshouse experiments, there will be trends along both axes and so designs with rows and columns have long been recommended for these experiments. Youden (1940) recommended the use of Latin and Youden squares and Cochran and Cox (1957, Section 4.3.1) recommended Latin square designs for experiments involving a single treatment factor. Edmondson (1989) used a Graeco-Latin square in a split plot design. Williams and John (1996) used factorial designs with rows and columns in designing glasshouse experiments and Williams et al. (2002, Section 7.5.1) advocated the use of designs with rows and columns for glasshouse experiments.

Tran (2009) reports a review of 20 ecological journals over the period from 1980 to 2006. The review focussed on articles concerning experiments in glasshouses or greenhouses on native plants that grow in temperate and semi-arid climates like Australia. In total, 59 experiments were reported, of which 43 involved factorial treatments. Only one of the 59 experiments stated that a design with rows and columns was used and this utilized a strip-plot design. This somewhat surprised us because our experience, and that of other statisticians, is that designs with rows and columns are often used. While this disparity might indicate that the designs being employed are not always correctly reported, it could also mean that a large proportion of ecological experiments are being designed by researchers themselves and that they do not use designs with rows and columns. Hence, while the evidence is not conclusive, it does seem that there is under-utilization of designs with rows and columns. Of course, they will not always be appropriate. However, to ensure they will be used whenever appropriate, every obstacle to their use needs to be removed. In this paper we facilitate the design of experiments for situations in which the number of rows does not equal the number of columns by extending the range of designs that can be readily constructed for this case.

Natural contenders for designs with rows and columns for factorial experiments are quasi-Latin designs. These designs were introduced by Yates (1937) for factorial experiments in which, like a Latin square, the treatments are to be applied to units arranged in an equal number of rows and columns. However, unlike Latin squares, the treatments are arranged such that no treatment occurs more than once in a row or a column and not all treatments occur in any row or column. A quasi-Latin square design may consist of one or more quasi-Latin squares and each quasi-Latin square contains one or more complete sets of treatments (Rao, 1946; Cochran and Cox, 1957). If there is more than one quasi-Latin square, the design is usually treated as a nested row-column design that is α-resolved (Shah, 1978), where α is the number of complete sets of treatments per square.

Quasi-Latin square designs extend the factorial designs that have treatment effects confounded with blocks to those that allow for two-way elimination of heterogeneity. They require the (partial) confounding of interactions with rows and columns within squares. Treatments must be equally replicated and the number of replicates is restricted. For example, consider a 2^3 factorial experiment. The eight treatments can be arranged in one or more 4×4 squares: the number of replicates for treatments must be a multiple of 2. We do not consider designs like that given by Cochran and Cox (1957, Plan 8.1b) to be quasi-Latin squares as they consist of Latin squares for subsets of the treatments; they do not fit into the class considered by Rao (1946) because the squares do not contain complete replicates and do not have treatment effects confounded within squares.

However, as has already been suggested, not all experiments in practice satisfy the restrictions placed on the number of replicates for a quasi-Latin square design and this was the case for experiments considered in Tran (2009). To provide more flexible designs we look to Latin rectangle designs. Adapting Preece (2006), Latin rectangles are defined to have k rows and ℓ columns for v treatments with $k \neq \ell$, $k \leq v$ and $\ell \leq v$. This differs from Preece (2006) in not insisting that $k < \ell$, although we will usually present designs so that this is true. Youden square designs, for which $\ell = v$, are Latin rectangles and are constructed from balanced incomplete block designs, with columns corresponding to blocks. Healy (1951) describes Latin rectangle designs for 2^k factorial experiments on a rectangle of 4×8 units. However, we refer to the subset of Latin rectangles in which a factorial set of treatments is assigned to a rectangle as *quasi-Latin rectangles*, because of their similarity to quasi-Latin squares. They retain the property of having no treatment repeated in any row or column. Then, a *quasi-Latin rectangle design* consists of one or more quasi-Latin rectangles, and may contain more than one complete replicate of the treatments. Thus, the designs given by Healy (1951) are described as quasi-Latin rectangle designs that assign 2^3, 2^4 or 2^5 factorial treatments to a single 4×8 rectangle. In addition, we consider *extended quasi-Latin rectangles* for which $k \neq \ell$ and at least one of k and ℓ exceeds v, so that either rows or columns or both contain some treatments more than once.

We begin in Section 2 by giving notation and some definitions, while Section 3 outlines some general principles that apply in designing experiments with rows and columns. In Section 4, a general method for the construction of row-column designs for symmetric factorial experiments is described and this is illustrated for a range of combinations of numbers of rows and columns in Section 5. Section 6 gives two further methods for parameter values satisfying certain conditions and Section 7 discusses the construction of row-column designs using the different methods and compares the designs obtained. The examples are only representative of the designs that can be constructed using the methods. The only type of design considered up to this point is row-column designs, the construction of designs with multiple squares or rectangles and choosing between these different types of design being deferred until Section 8. Some general aspects of quasi-Latin designs are discussed in Section 9.

2. Notation and some definitions

We consider designs in which there are s squares or rectangles each with k rows by ℓ columns, for $s \geq 1$ and $k, \ell \geq 2$. These squares and rectangles will be called *whole frames*. There are a total of v treatments, each with r replicates in each whole frame, and these treatments are the combinations of m factors each with p levels, where p divides both k and ℓ. Hence $v = p^m$ and $vr = k\ell$. Any subrectangle or subsquare of v units which contains one complete set of treatments will be called a *grid*. A single whole frame often contains grids of different shapes. The experimental unit, to which a single treatment is to be applied, is referred to simply as a *unit*. There are $sk\ell$ units in total. We assume that p is prime. This is not necessarily restrictive because, for a factor whose number of levels is a power of a prime, it is possible to substitute a combination of pseudofactors all of whose numbers of levels are equal to that prime.

Our construction methods use characters (Bailey, 2008, Section 12.2). Each level of a p-level factor is coded with the integers $0, 1, \ldots, p-1$. Each treatment combination of m factors can be written as an m-tuple of these levels. A character specifies a linear combination of factors that can be evaluated for each treatment combination; the coefficients are integers modulo p, as is the evaluation. For example, for the 3-level factors A and B, the levels are coded 0, 1 and 2 and one of their nine treatment combinations is $(2, 1)$. One character is $A + 2B$ and, for $(2, 1)$, it evaluates to $1 \times 2 + 2 \times 1 = 2 + 2 = 1$.

For sources we use the notation of Brien et al. (2011). In particular, A#B denotes the interaction of factors A and B, R [Q] denotes the nested effects of factor R within the levels of factor Q, and R [P ∧ Q] denotes the nested effects of R within the combinations of the levels of factors P and Q. For most designs in this paper, the only factors indexing the units before treatments are allocated are Rows and Columns, which are crossed. Hence, the unit sources are Rows, Columns and Rows#Columns.

3. Some principles in designing experiments with rows and columns

Firstly, as will be seen in this paper, often there are competing designs with different properties for fixed basic design parameters, such as the numbers of treatments, replicates, rows and columns. They may differ in the unit sources that are taken into account in the design and the manner in which the different treatment sources are confounded with the unit sources. Hence, choosing between designs depends on the expected sources of variability and potential treatment effects. With regard to the potential treatment effects, the issue is usually about which interactions, if any, need to be allowed for. If the designer decides that certain interactions are likely, then in view of the likely smaller size of interactions (Yates, 1937), it is especially important to maximize the amount of information about them which is confounded with the smallest source of variability. Consequently, our objective is not to obtain the "best design" for a given set of factorial treatments and of units, but to give several designs each of which is applicable in different situations.

Quasi-Latin designs are resolvable so that there is the option of developing a design that is (i) a *row-column design* with a single whole frame, (ii) *a nested design* being an α-resolved design consisting of s whole frames within which rows and columns are nested, or (iii) *a contiguous design*, which is like a nested design except that the contiguity between frames is acknowledged, so that treatments may be latinized to rows or columns (Williams, 1986). The value of α depends on the choice of s. Latinizing the treatments means that they are replicated as equally as possible in the direction being latinized. The three types of design that have been described differ in the sources of unit variability for which they allow. Hence, they have different unit structures and so different randomizations. A *unit structure* for an experiment is the decomposition of its data vector according to unit sources only, all treatment sources being disregarded.

A row-column design anticipates differences between rows and between columns. In this case, the factors indexing the units are Rows and Columns and these are crossed, as in, say, the design in Section 5.1. The randomization that applies is that rows and columns are permuted independently.

A nested design is appropriate when (i) there is a set of frames among which differences are anticipated and (ii) differences are also anticipated between rows and columns within each frame, these not being consistent across frames. Its factors are Frames, Rows and Columns, with Rows and Columns nested within Frames, as in Design 1 in Section 8.1, where the word Squares is used for Frames. For the randomization of a nested design, frames are permuted, as are rows and columns within each frame.

A contiguous design, like a nested design, has frames. It is utilized when, in addition to the unit variability for the nested design, consistent differences between rows, for horizontally-aligned frames, and columns, for vertically-aligned frames, are expected across frames. To account for these consistent differences, the treatments are latinized across frames to rows or columns or rows and columns, depending on the contiguity of the design. For designs in which either rows or columns are contiguous, the factors are Frames, Rows and Columns. If rows are contiguous, because consistent difference between rows across frames are expected, then Rows are crossed with Frames and Columns, and Columns are nested within Frames. An example is Design 3 in Section 8.1, where the word Squares is used for Frames. Randomization involves the permutation of frames, of rows, and of columns within frames.

Yates (1937) originally stressed that all rows and all columns should be randomized, because he was concerned to ensure that the Latin square analysis would be unbiased. However, the other randomizations can be applied, provided the appropriate analysis is then employed. The choice between the three options should be based on the expected unit sources of variability, rather than on the availability of a resolvable design. Will there be differences between frames? Will rows or columns differ consistently across frames? If it is likely that there are differences between frames and a row-column design is used, then either the differences between frames may obscure differences between treatments or, depending on the outcome of the randomization, the Residual mean square will give an overestimate of the variance of treatment effects. Similarly, if a nested design is used when a contiguous one is needed, then consistent differences between rows across frames will inflate either the Residual mean square for Rows within Frames or the apparent size of treatment effects confounded with Rows within Frames. On the other hand, if such differences are not appreciable, there is a penalty in employing a nested or contiguous design in that they are often less efficient than a good row-column design. Especially, for smaller designs, the loss in efficiency may outweigh any advantage of a nested or contiguous design.

Some authors advocate the use of a nested design, or, if frames are contiguous, a contiguous design in which treatments are latinized, whenever feasible and suggest that terms for the extra required sources of variation be omitted if a preliminary data analysis indicates that they are minor sources. This amounts to a "sometimes-pool" strategy. Janky (2000), in reviewing this topic, concluded that this strategy should not be used routinely as it generally inflates the probability of Type I errors and offers, at best, insubstantial gains in power. In a similar vein, Gilmour and Goos (2009) warn of the dangers that arise from omitting variance components that happen to have zero estimates and are based on small degrees of freedom. Our conclusion is that the designer should select a design appropriate for the anticipated sources of variation and use the analysis appropriate for the chosen design. This supports our objective of having a range of designs for a specific set of design parameters.

In considering the confounding of factorial treatment sources with unit sources, the *canonical efficiency factor* will be used because it reflects the amount of information about a treatment source, adjusted for previously fitted treatment sources, that is (partially) confounded with a particular unit source (John and Williams, 1995, Section 2.3). One property that distinguishes between designs is whether or not they are structure balanced (Brien and Bailey, 2009). This obtains when, for all treatment sources (partially) confounded with a particular unit source, (i) each treatment source has a single canonical efficiency factor and (ii) treatment sources remain orthogonal when estimated from that unit source. Condition (i) is met for all 2^m factorial experiments. Condition (ii) means that the experiment has *orthogonal factorial structure* (Bailey, 1985). The advantage of factorial experiments having structure balance is that the estimates of various factorial effects are independent. It is desirable that, if at all possible, a structure-balanced design is used, even if it means sacrificing some efficiency on particular treatment effects. A structure-balanced design has a much simpler analysis, and conclusions are more straightforward than for designs that are not structure balanced. Also, the amount of treatment information available from unit sources is maximized for structure-balanced experiments, the amount of information available about treatments from a unit source being the weighted sum of the degrees of freedom of treatment sources confounded with it, the weights being the efficiency factors.

We have obtained designs in two distinctly different ways: (i) choosing which treatment effects to confound with rows and columns and constructing the designs using the methods introduced in Sections 4 and 6, and (ii) using a package for the computer generation of designs based on an interchange algorithm, in our case CycDesigN (CycSoftware Ltd, 2009). The first gives the designer greater control over the design by providing the tools for choosing how to spread the information about the treatment effects across the different sources of variability, that is, deciding what to confound a treatment effect with and the amount of information that will be associated with more variable unit sources.

On the other hand, software, such as CycDesigN, has the advantage of giving an automated procedure, which requires little more than input of the design parameters. CycDesigN 4.0 (CycSoftware Ltd, 2009), optimizes the overall weighted efficiency of a design. A further significant advantage of such programs is that they can produce designs for a wider range of design parameters than the construction methods we outline. With CycDesigN, some control over the properties of the design produced is afforded by varying the weighting of (i) treatment main effects relative to two-factor treatment interactions and (ii) the row, column and row-column components of a resolved design relative to each other. For (i), the weighted efficiency is the weighted linear combination of the average efficiency factors of the treatment main effects and of the treatment two-factor interactions (CycSoftware Ltd, 2009, Section 9.3). Hence, it is not possible to specify the properties of single main effects or two-factor interactions or anything about the higher order interactions. Also, CycDesigN generates only 1-resolved designs.

It is our contention that, in spite of the availability of software like CycDesigN, construction methods of the type outlined in this paper are useful because they permit more direct control of the design process and can lead to designs that are better suited to some particular situations than those produced by such software. Further, they often have nice properties and provide a point of comparison for computer-generated designs. The methods of construction will be useful to any designer who does not have access to software like CycDesigN, provided the experimental conditions correspond to the parameter combinations available with the methods. In comparing constructed and computer-generated designs, we generally compare the canonical efficiency factors of the two designs. However, even if the canonical efficiency factors are the same, the designs themselves may not be isomorphic in that it is not possible to obtain one by permuting the other in ways that are allowable for its unit structure.

4. A method for the construction of row-column designs for symmetric factorial experiments

In this section we present Method 1 for constructing whole frames of shape $k \times \ell$. We assume that $v = p^m$, that v divides $k\ell$, that k and ℓ are both divisible by p. The replication is r, where $r = k\ell/v$. The method produces quasi-Latin squares and rectangles, extended quasi-Latin rectangles and Latin squares. It involves dividing a whole frame into several types of frames as illustrated in Figure 1 and consists of the steps below. The crux of the method is to form what we term box frames, whose dimensions are powers of p such that each contains one or more complete replicates of the treatments. Then sets of characters can be confounded with sets of rows and sets of columns in each box frame.

Step 1 Divide up the whole frame: Having ascertained the values of p, m, k, ℓ and r, determine values of t and u so that we can write $k = p^t r_1$ and $\ell = p^u r_2$, where $1 \leq t \leq m$, $1 \leq u \leq m$, $t + u \geq m$. That is, we factorize both k and ℓ as a product of two integers, of which the first is p raised to a non-zero power. The condition $t + u \geq m$ means that the product of the powers of p must be divisible by v. Select t and u as follows:

(a) If v divides k, then take $t = m$ and $r_1 = k/v$; if k is a power of p smaller than v

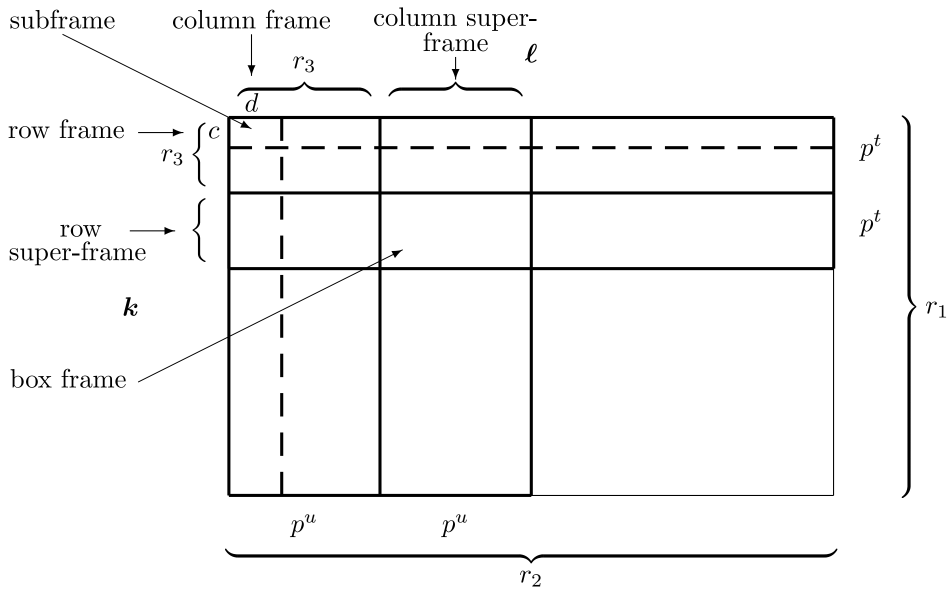

Fig. 1. Division of the whole frame for Method 1: the whole frame is divided into r_1 row super-frames and r_2 column super-frames whose intersections form box frames of shape $p^t \times p^u$; each row super-frame is divided into r_3 row frames and each column super-frame is divided into r_3 column frames; their intersections form subframes of shape $c \times d$.

or r is not divisible by p, then p^t is the largest power of p dividing k; otherwise there is some choice in the value of t.

(b) If v divides ℓ, then take $u = m$ and $r_2 = \ell/v$; if ℓ is a power of p smaller than v or r is not divisible by p, then p^u is the largest power of p dividing ℓ; otherwise there is some choice in the value of u.

Note that if $k < v$, $\ell < v$ and r is a power of p, then $r_1 = r_2 = 1$.

Now divide the whole frame into r_1 *row super-frames* of p^t whole rows and r_2 *column super-frames* of p^u whole columns. The intersection of a row super-frame and a column super-frame forms a *box frame* of shape $p^t \times p^u$. Of course, if $r_1 = r_2 = 1$, then the super-frames and box frames are all the same as the whole frame.

To set up row and column frames, calculate $c = p^{m-u}$, $d = p^{m-t}$ and $r_3 = p^{t+u-m}$. Then $r = r_1 r_2 r_3$. Divide each row super-frame into r_3 row frames of shape $c \times \ell$ and each column super-frame into r_3 column frames of shape $k \times d$. The intersection of a row frame and a column frame forms a subframe of shape $c \times d$ and each box frame contains an $r_3 \times r_3$ array of these subframes. Also, each box frame contains r_3 grids of shape $p^t \times d$, as well as r_3 of shape $c \times p^u$.

Step 2 Specify the row design: Each row frame consists of r_2 grids of shape $c \times p^u$. If $u = m$ then $c = 1$. In this case, each $1 \times p^u$ grid contains a complete replicate of the treatments and there is no need to further consider the row design. If $c > 1$, then select $r_1 r_3$ sets of characters, one set per row frame, so that each set specifies $(c - 1)$ treatment degrees of freedom to confound with c rows. The characters specifying one lot of $(c - 1)$ treatment degrees of freedom must be closed under the formation of sums (modulo p). We shall call these *row characters*. Each set divides the treatments into c groups of size p^u. If $r_2 = 1$ then the groups in each row frame are completely confounded with rows. If $r_2 > 1$ then each row frame needs a $c \times r_2$ row-column design Δ_1 for c treatments as an auxiliary design, where these treatments correspond to the c groups of treatments defined by the set of row characters for this row frame. In design Δ_1, the columns are complete and the row design should be as efficient as possible.

Step 3 Specify the column design: Similarly, each column frame consists of r_1 grids of shape $p^t \times d$. If $t = m$ then $d = 1$. In this case, each $p^t \times 1$ grid contains a complete replicate of the treatments and there is no need to further consider the column design. If $d > 1$, then select $r_2 r_3$ sets of characters, each specifying $d - 1$ degrees of freedom and dividing the treatments into d groups of size p^t. We shall call these *column characters*. If $r_1 > 1$ then each column frame needs a $r_1 \times d$ auxiliary design Δ_2 for d treatments.

Step 4 Ensure a unique treatment for each unit: In each box frame, the treatments in each $1 \times p^u$ subrectangle are specified: if $u = m$, this subrectangle contains a complete set of treatments; otherwise they are specified by the row characters and, if $r_2 > 1$, the auxiliary design Δ_1. Similarly, if $t = m$, then each $p^t \times 1$ subrectangle contains a complete set of treatments; otherwise the treatments it contains are specified by the column characters and, if $r_1 > 1$, the auxiliary design Δ_2. If $r_3 = 1$, this uniquely determines the treatment on each unit. Otherwise, for each box frame, choose a set of characters which divide the treatments into r_3 groups of size cd. We shall call these *unit characters*. The groups are assigned to the $r_3 \times r_3$ array of subframes of shape $c \times d$ by a using a $r_3 \times r_3$ Latin square Δ_3 as third auxiliary design. For each box frame, the sets of characters of whichever of the three different types (row, column and unit) are needed must satisfy the following condition:

any nonempty collection of characters, all of different types, must be linearly independent modulo p. (1)

If $t = u = m$, then $c = d = 1$ and $r_3 = p^m$ so that no row and column characters are required and there is no need to specify unit characters. All that is needed is Δ_3, which is a $p^m \times p^m$ Latin square. The whole design is an $r_1 \times r_2$ array of such Latin squares.

In general, for each set of c rows, one has to specify either (i) c characters, including 0, closed under addition, or (ii) $m-u$ linearly independent characters, or (iii) $(c-1)/(p-1)$ characters none of which is a multiple of any other. Similarly, for each set of d columns, one has to specify either (i) d characters, including 0, closed under addition, or (ii) $m-t$ linearly independent characters, or (iii) $(d-1)/(p-1)$ characters none of which is a multiple of any other. A set of characters to be confounded with c rows (d columns) can be repeated amongst the r_1r_3 (r_2r_3) sets of row (column) characters. If the sets of one type are not all the same, this results in partial confounding.

If r is divisible by p but is not a power of p then there is some choice in the values of u and t. Different choices may lead to designs with different properties. If $t+u=m$ then $r_3=1$ and there is no need for unit characters, so Condition (1) is easier to satisfy. On the other hand, there is more freedom of choice for the row characters when c is smaller, and more freedom of choice for the column characters when d is smaller. The availability of good $c \times r_2$ and $r_1 \times d$ row-column designs for the possible values of c, d, r_1 and r_2 is also an issue. When u is larger then c and r_2 are smaller so the former are easier to find, but there may be more choice when c and r_2 are larger. For designs of practical size, it seems unlikely that all three of r_1, r_2 and r_3 will be bigger than one.

With $k=\ell$ and $r_1=r_2=1$, the method is equivalent to that of Rao (1946) for constructing quasi-Latin square designs. That is, our method generalizes that of Rao (1946) in two ways. The first simply allows $t \neq u$ when $r_1=r_2=1$. The second allows one or both of r_1 and r_2 to be bigger than one: in either case, another auxiliary design is needed.

5. Examples of quasi-Latin squares and rectangles with dimensions less than the number of treatments

5.1. A 2^3 factorial in a 4×4 square

For $p=2$, $m=3$, $k=\ell=4$ and $r=2$, both k and ℓ are powers of p and the only possible quasi-Latin square is for $t=u=2$ and $r_1=r_2=1$. Thence, $c=d=2$ and $r_3=2$. Hence, we can ignore super-frames and take the box frame to be the whole frame. The square is subdivided into two row frames of shape 2×4 and two column frames of shape 4×2. The whole (box) frame is a 2×2 array of subframes, each of which is 2×2.

Construction of the design requires two row characters and two column characters, as well as a unit character; the unit character splits the 8 treatments into 2 groups of 4 and the groups are assigned using, as an auxiliary design Δ_3, a 2×2 Latin square.

Let U and V be the row characters, W and X be the column characters and Y be the unit character. They can be any five characters with the property that none of $U+W$,

Table 1. Quasi-Latin square for a 2^3 factorial experiment in 4 rows × 4 columns

	$A+B$		$A+C$	
	$=0$	$=1$	$=0$	$=1$
$B+C=0$	$1,1,1$	$1,0,0$	$0,0,0$	$0,1,1$
$B+C=1$	$1,1,0$	$1,0,1$	$0,1,0$	$0,0,1$
$A+B+C=0$	$0,0,0$	$0,1,1$	$1,0,1$	$1,1,0$
$A+B+C=1$	$0,0,1$	$0,1,0$	$1,1,1$	$1,0,0$

Table 2. Canonical efficiency factors and Residual degrees of freedom (DF) for a 2^3 factorial experiment in 4 rows × 4 columns

	Treatment sources							*Residual*
Unit sources	A	B	C	A#B	A#C	B#C	A#B#C	*DF*
Rows	0	0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$	1
Columns	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0	0	1
Rows#Columns	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	2

$U+X$, $V+W$ and $V+X$ is equal to Y or to 0. The four rows are defined by $U=0$, $U=1$, $V=0$ and $V=1$, respectively, and the four columns by $W=0$, $W=1$, $X=0$ and $X=1$, respectively. These restrictions are not enough to define the entries uniquely, so we put $Y=1$ on the top left-hand and the bottom right-hand subsquares, and put $Y=0$ on the other two subsquares. In the top left-hand corner, the four combinations of levels of U and W, together with the constraint $Y=1$, define the treatments uniquely, giving all four treatments with $Y=1$. Similarly, in the top right-hand corner, the four combinations of levels of U and X, together with the constraint $Y=0$, define the treatments uniquely, giving the remaining four treatments. Hence the first two rows form a complete replicate. In a similar manner, the treatment on each unit is defined uniquely, and the first two columns form a complete replicate, as do the last two rows and also the last two columns.

This construction results in U and V each losing half their information to rows, if $U \neq V$, while W and X each lose half their information to columns, if $W \neq X$. The character Y is necessary for the construction, but it remains orthogonal to both rows and columns.

For example, if we want full information on all main effects then we can put $U=B+C$, $V=A+B+C$, $W=A+B$, $X=A+C$ and $Y=A$. This gives the design in Table 1. Up to relabelling of the factors, it is the same as Square II in Cochran and Cox (1957, Table 8.1). The canonical efficiency factors and Residual degrees of freedom for the design are in Table 2. Clearly, the design has too few Residual degrees of freedom to be of practical use.

To increase the Residual degrees of freedom, two squares ($s=2$) are usually proposed for a 2^3 factorial. Cochran and Cox (1957) give such a plan, which will be compared with other designs using two squares in Section 8.1. However, there is another possibility that applies when the rows (or columns) of the two squares are contiguous. Namely, construct a single 4×8 rectangle, as is done in Section 7.1.

5.2. A 2^5 factorial in an 8×8 square

Here $p = 2$, $m = 5$, $k = \ell = 8$ and $r = 2$ and so both k and ℓ are powers of p. Hence, the only possible quasi-Latin square is for $t = u = 3$ and $r_1 = r_2 = 1$ so that $c = d = 4$ and $r_3 = 2$. Again, super-frames are superfluous. The row frames are 4×8 and column frames are 8×4 and there are two of each. To construct the design two sets of three row characters are needed and two sets of three column characters. Plan 8.3 in Cochran and Cox (1957) uses the two sets $\{A + B + C, A + D + E, B + C + D + E\}$ and $\{A + B + D, B + C + E, A + C + D + E\}$ for row characters and the two sets $\{A + C + E, B + C + D, A + B + D + E\}$ and $\{A + C + D, B + D + E, A + B + C + E\}$ for column characters. Each set of characters is closed under addition (modulo 2). The box frame for this design is of shape 8×8 or the whole frame. As $r_3 = 2$, the whole frame consists of a 2×2 array of subframes of shape 4×4 and a unit character is required. The unit character chosen is $A + B + C + D$ and a 2×2 Latin square is used to assign its levels to the subframes.

In Section 8.2, nested and contiguous designs, based on two 4×8 grids, are explored as alternatives to the above design.

5.3. A 3^3 factorial experiment in a 9×12 rectangle

In this case $p = 3$, $m = 3$, $k = 9$, $\ell = 12$ and $r = 4$ so that k is a power of p and r is not divisible by p. Thus, $t = 2$ and $u = 1$. We have $c = 9$, $d = 3$ and $r_3 = 1$. The numbers of row and column super-frames are $r_1 = 1$ and $r_2 = 4$. Then there is one row frame, the same as the row super-frame and the whole frame; each column super-frame is also a column frame and a box frame, and consists of a 9×3 grid.

Row characters specifying 8 degrees of freedom and four column characters, each specifying 2 degrees of freedom, are required. An auxiliary design Δ_1, for assigning the nine groups defined by the row characters, is needed and this will be a 9×4 row-column design for 9 treatments. A suitable design has the following rows: $(5, 6, 8, 9)$, $(9, 4, 6, 7)$, $(7, 8, 4, 5)$, $(8, 9, 2, 3)$, $(3, 7, 9, 1)$, $(1, 2, 7, 8)$, $(2, 3, 5, 6)$, $(6, 1, 3, 4)$, $(4, 5, 1, 2)$. Use two of the row characters to index treatments 1–9 in lexicographical order. Then, the four degrees of freedom corresponding to these two row characters have canonical efficiency factor $1/4$ in rows, while the canonical efficiency factor for the four degrees of freedom for the other two row characters is $1/16$.

No unit characters are required because $r_3 = 1$.

For example, one could choose $A+B$ and $B+C$ for row characters, so that $A+2C$ and $A+2B+C$ would be required to make the complete set of row characters. The column characters could be chosen from $A+B+C$, $A+B+2C$ and $A+2B+2C$. For example, one could use two of these characters in one frame each and the other in two frames, thus partially confounding the corresponding effects. Those used in just one would have 75% of their information orthogonal to rows and columns and, for the other character, it would be 50%.

5.4. *A* 2^4 *factorial in an* 8×12 *rectangle*

In this example $p = 2$, $m = 4$, $k = 8$, $\ell = 12$ and $r = 6$. We have $k = 2^3$ and $\ell = 2^2 3^1$ and so, as k is a power of p, it must be that $t = 3$ and, as r is divisible by p, there is a choice of values for u; $u = 2$ is chosen. As a result $c = 4$, $d = 2$ and $r_3 = 2$. The numbers of row and column super-frames are $r_1 = 1$ and $r_2 = 3$, respectively. Hence, there are two row frames of shape 4×12 in the one row super-frame, and three column frames, one in each column super-frame of shape 8×4.

The row and column characters chosen for this design are given in Table 3. Because $r_2 = 3$, an auxiliary design Δ_1 is needed to assign groups defined by the row characters to the 4×3 array of subrectangles of shape 1×4. The transpose of a 3×4 Youden square, constructed by removing the last row from a Latin square, is suitable. The three rows of the Youden square are $(1, 2, 3, 4)$, $(2, 3, 4, 1)$ and $(3, 4, 1, 2)$. As the Youden square has $1/9$ of the treatment information confounded with columns, $1/9$ of each of the row characters is confounded with Rows. Because $r_1 = 1$, an auxiliary design Δ_2 is not needed for assigning the column characters.

The box frames for this design are of shape 8×4. As $r_3 = 2$, a box frame consists of a 2×2 array of subframes of shape 4×2 and a unit character is required for each box frame. They are in Table 3. The auxiliary design Δ_3, used in assigning unit characters, is a 2×2 Latin square with rows (0,1) and (1,0).

The canonical efficiency factors and Residual degrees of freedom for this design are summarized in Table 4. This shows that the design has very good properties. Many other choices of sets of confounding characters are possible, depending on which interactions are considered important.

5.5. *A* 2^3 *factorial in a* 6×12 *rectangle*

Here $p = 2$, $m = 3$, $k = 6$ and $\ell = 12$, so that $r = 9$ and, as r is not divisible by p, we are forced to put $t = 1$, $u = 2$ and $r_1 = r_2 = 3$, which give $c = 2$, $d = 4$ and $r_3 = 1$. Thus the row and column super-frames are the same as the row and column frames.

There are three row frames, each of shape 2×12. We can assign the characters A, B and

Table 3. Quasi-Latin rectangle for a 2^4 factorial experiment in 8 rows $\times$ 12 columns

	Column frame I				*Column frame II*				*Column frame III*			
Unit characters	A				D				$A+B+C+D$			
Column characters	$A+B+C+D$		$A+C+D$		$A+B+C$		$C+D$		$A+B+D$		$B+C+D$	
	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$
Upper row frame	0,0,0,0	0,0,0,1	1,1,1,0	1,1,1,1	1,1,0,0	0,0,1,0	0,0,1,1	1,1,0,1	0,1,0,1	1,0,1,0	1,0,1,1	0,1,0,0
(Row characters:	0,0,1,1	0,0,1,0	1,1,0,1	1,1,0,0	1,0,1,0	0,1,0,0	1,0,1,1	0,1,0,1	1,0,0,1	0,1,1,0	1,0,0,0	0,1,1,1
$A+B$, $A+C$, $B+C$)	0,1,0,1	0,1,0,0	1,0,1,0	1,0,1,1	0,1,1,0	1,0,0,0	0,1,1,1	1,0,0,1	0,0,0,0	1,1,1,1	1,1,1,0	0,0,0,1
	0,1,1,0	0,1,1,1	1,0,0,1	1,0,0,0	0,0,0,0	1,1,1,0	1,1,1,1	0,0,0,1	1,1,0,0	0,0,1,1	1,1,0,1	0,0,1,0
Lower row frame	1,1,1,1	1,1,0,1	0,0,0,0	0,0,1,0	0,0,0,1	0,0,1,1	1,1,0,0	1,1,1,0	1,0,1,1	0,1,0,0	0,1,1,0	1,0,0,1
(Row characters:	1,1,0,0	1,1,1,0	0,0,1,1	0,0,0,1	1,0,1,1	1,0,0,1	0,1,0,0	0,1,1,0	0,1,1,1	1,0,0,0	0,1,0,1	1,0,1,0
$A+D$, $B+D$, $A+B$)	1,0,0,1	1,0,1,1	0,1,0,0	0,1,1,0	0,1,1,1	0,1,0,1	1,0,0,0	1,0,1,0	0,0,1,0	1,1,0,1	0,0,0,0	1,1,1,1
	1,0,1,0	1,0,0,0	0,1,1,1	0,1,0,1	1,1,0,1	1,1,1,1	0,0,0,0	0,0,1,0	1,1,1,0	0,0,0,1	0,0,1,1	1,1,0,0

Table 4. Canonical efficiency factors and Residual degrees of freedom (DF) for the design for a 2^4 factorial experiment in 8 rows × 12 columns

	Unit sources		
Treatment sources	Rows	Columns	Rows#Columns
A, B, C, D	0	0	1
A#B	$\frac{1}{9}$	0	$\frac{8}{9}$
A#C, A#D, B#C, B#D	$\frac{1}{18}$	0	$\frac{17}{18}$
C#D, A#B#C, A#B#D, A#C#D, B#C#D, A#B#C#D	0	$\frac{1}{6}$	$\frac{5}{6}$
Residual DF	2	5	62

Table 5. Quasi-Latin rectangle for a 2^3 factorial experiment in 6 rows × 12 columns

	Column frame I				*Column frame II*				*Column frame III*			
$A = 0,0,1$	0,0,0	0,0,1	0,1,0	0,1,1	0,0,0	0,0,1	0,1,0	0,1,1	1,1,1	1,1,0	1,0,1	1,0,0
$A = 1,1,0$	1,1,1	1,1,0	1,0,1	1,0,0	1,1,1	1,1,0	1,0,1	1,0,0	0,0,0	0,0,1	0,1,0	0,1,1
$B = 0,0,1$	0,0,1	1,0,1	1,0,0	0,0,0	0,0,1	1,0,1	1,0,0	0,0,0	1,1,0	0,1,0	0,1,1	1,1,1
$B = 1,1,0$	1,1,0	0,1,0	0,1,1	1,1,1	1,1,0	0,1,0	0,1,1	1,1,1	0,0,1	1,0,1	1,0,0	0,0,0
$C = 0,0,1$	0,1,0	1,0,0	0,0,0	1,1,0	0,1,0	1,0,0	0,0,0	1,1,0	1,0,1	0,1,1	1,1,1	0,0,1
$C = 1,1,0$	1,0,1	0,1,1	1,1,1	0,0,1	1,0,1	0,1,1	1,1,1	0,0,1	0,1,0	1,0,0	0,0,0	1,1,0

C to one row frame each. In each row frame we use, for the two levels of the row character, the 2×3 auxiliary design Δ_1 whose rows are $(0,0,1)$ and $(1,1,0)$: this confounds $1/9$ of the between-level information with rows. There are three column frames, each of shape 6×4: in order to satisfy Condition (1), we take $\{A+B, A+C, B+C\}$ to be the set of column characters in each column frame. This set divides the eight treatments into four groups of two, so our 3×4 auxiliary design Δ_2 is the Youden square given in Section 5.4. There is no need for a unit character or a third auxiliary design, because $r_3 = 1$.

The complete design is shown in Table 5. All main effects have canonical efficiency factors $1/27$, 0 and $26/27$ in Rows, Columns and Rows#Columns respectively, while the corresponding figures for the two-factor interactions are 0, $1/9$ and $8/9$. The three-factor interaction is completely confounded with Rows#Columns.

6. Other methods for constructing row-column designs for symmetric factorial experiments

We now give two other methods for constructing (extended) quasi-Latin rectangles. Method 2 applies when one of k and ℓ is a multiple of v. Method 3 divides the design into unequally-sized segments and a design is constructed for each segment.

For Method 2, not only must one of k and ℓ be a multiple of v, but the other must be a proper divisor of v. Take ℓ to be a multiple of v; for the case of k a multiple of v interchange the roles of rows and columns. While such designs can be constructed using Method 1, this requires the specification of both column and unit characters. On the other hand, Method 2 requires only column characters and so with it there will usually be more choice for the column characters. Further, Condition (1) is vacuously satisfied, and so there are no constraints on the choice of column characters. Hence, Method 2 is likely to be the preferred method for this class of designs, unless the designer is prepared to confound column characters with multiple column frames. The steps for Method 2 are:

Step 1: Divide up the whole frame: Divide the design into r_2 column super-frames of shape $k \times v$, where $r_2 = \ell/v$. Divide each column super-frame into k column frames of shape $k \times d$, where $d = v/k$.

Step 2: Specify the column design: In each column super-frame, choose k sets of column characters each specifying $d - 1$ degrees of freedom. It is not necessary for all the sets to be different. Each set of characters is confounded with the columns of one of the column frames.

Step 3: Form the row design: In each column super-frame, rearrange the treatments in each column, using the algorithm given in Technique 11.1 of Bailey (2008), so that each row consists of a complete replicate.

The justification for the last step is that the column design can be viewed as a symmetric incomplete-block design. By Hall's Marriage Theorem, the treatments in each column can be rearranged so that each row consists of a complete replicate.

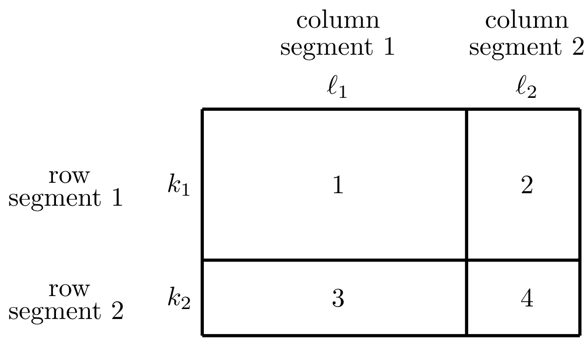

Fig. 2. Segmentation of the whole frame for Method 3 into four segments numbered as shown

Method 3 divides the design into segments as illustrated in Figure 2. It is useful when at least one of k and ℓ is neither a power of p nor a multiple of v; otherwise, it duplicates

Method 1 or Method 2. W assume that the normal conditions for (extended) quasi-Latin designs apply.

Step 0: Initialize: Set $k_1 = k$, $k_2 = 0$, $\ell_1 = \ell$ and $\ell_2 = 0$.

Step 1: Choose the row segment sizes: If k is neither a power of p nor a multiple of v, then choose a value of t such that $p^t < k$, $t \leq m$, v divides $p^t\ell$, and p^t does not divide k. If there is no such value of t, then there is nothing to be gained by row segmentation. Otherwise, it is usually sensible to choose the largest possible value of t; in particular, if $k \geq v$ then take $t = m$. Let k_1 be the largest multiple of p^t which is smaller than k, and put $k_2 = k - k_1$. Then v divides $k_1\ell$ and $k_2\ell$, and p divides k_1 and k_2.

Step 2: Choose the column segment sizes: If ℓ is neither a power of p nor a multiple of v, then choose a value of u such that $p^u < \ell$, $u \leq m$, v divides $p^u k$, and p^u does not divide ℓ. If there is no such value of u, then there is nothing to be gained by column segmentation. Otherwise, it is usually sensible to choose the largest possible value of u; in particular, if $\ell \geq v$ then take $u = m$. Let ℓ_1 be the largest multiple of p^u which is smaller than ℓ, and put $\ell_2 = \ell - \ell_1$. Since v divides $p^u k$, we must have k divisible by p^{m-u}. If t is defined, then $m - u \leq t$, and hence p^{m-u} divides k_1 and k_2; otherwise $k_1 = k$ and $k_2 = 0$ and again p^{m-u} divides k_1 and k_2. Therefore v divides $k_1\ell_1$, $k_1\ell_2$, $k_2\ell_1$ and $k_2\ell_2$, and p divides ℓ_1 and ℓ_2.

Step 3: Divide the whole frame into segments: If $k_2 = \ell_2 = 0$, then it is not useful to segment the design and this method does not apply. Otherwise, segment the design as shown in Figure 2. Only if both $k_2 \neq 0$ and $\ell_2 \neq 0$ will there be four segments. If only one is nonzero, then there will be two segments.

Step 4: Construct a design for each segment: Use Method 1, 2 or 3, as appropriate, on each of the design segments. If there are two segments in the same row segment and row characters are needed for both, then, to minimize the amount of information on row characters in Rows in the whole design, the row characters in each row frame of the first segment should be a subset of those in the corresponding row frame of the second segment; this will also require the values of these characters in the corresponding rows of the two designs to be chosen suitably. Similar considerations apply if there are two segments in the same column segment and column characters are needed for both. The simplest situation is that $k_1 = \ell_1 = v$ and so a Latin square can be used for segment 1. In this situation, segment 2 will require only row characters for its construction, segment 3 will require only column characters and segment 4 will require both row and column characters, but these can be chosen independently of the those used for the other segments.

7. Examples comparing (extended) quasi-Latin rectangles constructed using the different methods

In this Section we compare several row-column designs for three sets of basic design parameters using all three methods of construction that we have presented, as well as computer search.

7.1. A 2^3 factorial in a 4×8 rectangle

For this example $p = 2$, $m = 3$, $k = 2^2$, $\ell = 2^3$ and $r = 4$. Here, we compare the properties of three designs, all of which have orthogonal factorial structure.

Design 1: As $\ell = v$, Method 2 in Section 6 applies with $r_2 = 1$ and $d = 2$. We use it to construct a design. There are four column frames of shape 4×2. A column character is needed for each column pair. For example, assign each of $A+B$, $A+C$, $B+C$ and $A+B+C$ to be confounded with one pair of columns, so that a different character is used for each pair. Then, 1/4 of the information on each of the corresponding effects is lost to columns. Table 6 shows one of the possible designs obtained after rearranging treatments in each column to make each row a complete replicate.

Design 2: The optimal design produced by CycDesigN 4.0, which can be constructed using Method 2 by assigning $A+C$ to one pair of columns and $A+B+C$ to three pairs of columns. It partially confounds 1/4 of A#C and 3/4 of A#B#C with its columns.

Design 3: Healy's (1951) design, which can be constructed using Method 1, with $r_1 = r_2 = 1$, $t = 2$ and $u = 3$ so that $c = 1$ and $d = 2$. Thus, $r_3 = 4$ and the whole (box) frame consists of a 4×4 array of 1×2 subframes. Four column characters are needed, as well as one set of three unit characters that are closed under addition and an auxiliary design Δ_3, for assigning the four groups of treatments determined by the four combinations of the values of the unit characters. Healy's design has the character $A + B + C$ assigned to every pair of columns, uses the set $\{B, C, B + C\}$ for unit characters, and takes a 4×4 Latin square for the auxiliary design. So, the interaction $A \# B \# C$ is totally confounded with columns and no treatment effects are confounded with rows.

Table 7 compares the canonical efficiency factors and Residual degrees of freedom for the three designs. In using CycDesigN to produce Design 2, the default weights ratio of $1 : 0.25$ for main effects relative to two-factor interactions was employed. One might consider reducing the weights for two-factor interactions in order to produce Design 1, which has lower efficiency for two-factor interactions. However, for two designs with orthogonal main effects and so maximal main-effect efficiency, such as Designs 1 and 2, the weighted efficiency

Table 6. Quasi-Latin rectangle for a 2^3 factorial experiment in 4 rows $\times$ 8 columns constructed using Method 2

$A+B$		$A+C$		$B+C$		$A+B+C$	
$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$
$0,0,0$	$1,0,0$	$0,1,0$	$0,0,1$	$0,1,1$	$1,1,0$	$1,0,1$	$1,1,1$
$1,1,0$	$1,0,1$	$0,0,0$	$1,0,0$	$1,1,1$	$0,0,1$	$0,1,1$	$0,1,0$
$0,0,1$	$0,1,0$	$1,1,1$	$0,1,1$	$0,0,0$	$1,0,1$	$1,1,0$	$1,0,0$
$1,1,1$	$0,1,1$	$1,0,1$	$1,1,0$	$1,0,0$	$0,1,0$	$0,0,0$	$0,0,1$

Table 7. Canonical efficiency factors and Residual degrees of freedom (DF) for the designs for a 2^3 factorial experiment in 4 rows $\times$ 8 columns

		Treatment sources							*Residual*
Design	*Unit sources*	A	B	C	A#B	A#C	B#C	A#B#C	*DF*
1	Rows	0	0	0	0	0	0	0	3
	Columns	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	3
	Rows#Columns	1	1	1	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	14
2	Rows	0	0	0	0	0	0	0	3
	Columns	0	0	0	0	$\frac{1}{4}$	0	$\frac{3}{4}$	5
	Rows#Columns	1	1	1	1	$\frac{3}{4}$	1	$\frac{1}{4}$	14
3	Rows	0	0	0	0	0	0	0	3
	Columns	0	0	0	0	0	0	1	6
	Rows#Columns	1	1	1	1	1	1	0	15

factor for a given set of weights must be greater for the design for which the sum of the efficiency factors for two-factor interactions is greater. Hence, Design 2, or any design whose two-factor efficiencies include both 0.75 and 1 and no other values, will always have higher efficiency than Design 1 and so Design 1 will only be selected as the optimal design if one manages to stop the iterative search procedure prematurely. This is more likely to be possible if a relatively very small weight is used for two-factor interactions, such as in a weights ratio of 1:0.001, because this will make the differences between the efficiency of the designs small (< 0.001).

Clearly, Design 3 suits experiments in which it is appropriate to confound the three-factor interaction with the likely more variable Columns, such as when this interaction is anticipated to be negligible. On the other hand, as concluded in Section 3, Design 1 will be preferred if a three-factor interaction is thought to be highly likely and one wants to estimate it with good precision.

7.2. A 2^3 factorial in a 4×6 rectangle

This example is for $p = 2$, $m = 3$, $k = 4$, $\ell = 6$ and $r = 3$. As always, Method 1 applies. Neither ℓ nor k are multiples of v and so Method 2 is not applicable. On the other hand, ℓ is nether a multiple of v nor a power of p, so that Method 3 can be used.

Three designs will be constructed, ordered according to the amount of information partially confounded with Rows#Columns: the amount for A#B#C decreases and that for the two-factor interactions increases. They demonstrate how the designer can influence the spread of the information about the treatment effects across the unit sources and show the flexibility of our construction methods.

Design 1: In this design Method 1 is used. As r is not a multiple of p, it follows that $t = 2$ and $u = 1$ so that $r_1 = 1$ and $r_2 = 3$. Also, $c = 4$, $d = 2$ and $r_3 = 1$. Hence, there are 3 column super-frames, each containing a single 4×2 grid that is both a box frame and a subframe. To construct the design requires, firstly, a set of row characters specifying 3 treatment degrees of freedom and an auxiliary design Δ_1 for assigning groups of treatments determined by the row characters. Secondly, one column character for each column super-frame is needed. Unit characters and the associated auxiliary design Δ_3 are not required.

Let $\{U, V, U + V\}$ be the set of row characters and $\{W, X, Y\}$ the set of column characters. It is not necessary for all the column characters to be different, but Condition (1) must be satisfied. The three row characters divide the eight treatments into four groups of two, say S_1, S_2, S_3 and S_4. The transpose of the 3×4 Youden square used in Section 5.4 is a suitable auxiliary design for assigning these groups.

To maximize the minimum canonical efficiency factor for all treatment effects when (partially) confounded with Rows# Columns, we can take $U = A$, $V = B$, $W = A+C$, $X = B + C$ and $Y = A + B + C$. Table 8 shows the final design.

Design 2: This design uses Method 3. Because k is a power of p, row segmentation is not useful and $k_2 = 0$. On the other hand, ℓ is not a power of p or a multiple of v, and $\ell > 4$ so that column segmentation can be employed. Here $u = 2$ so that $\ell_1 = 4$ and $\ell_2 = 2$. That is, segment 1 is of shape 4×4 and the other segment is 4×2. The first can be constructed as a 4×4 quasi-Latin square and the other as a 4×2 quasi-Latin rectangle, both using Method 1.

For the quasi-Latin rectangle, which consists of a single grid, a set of row characters specifying 3 treatment degrees of freedom and a column character are required. Suppose that, in order to have no main effects involved, the row characters are $A + B$, $A+C$ and $B+C$ and the column character is $A+B+C$. The row characters divide the treatments into four groups of two, one for each combination of the values of $A + B$, $A + C$.

For the quasi-Latin square, which has the same basic design parameters as the design given in Section 5.1, two row and two column characters, as well as a unit character, are needed. To match the quasi-Latin rectangle, the row characters for the quasi-Latin square should be a subset of those for the rectangle. Take $A + B$ and $A + C$. For the column characters, again to have no main effects involved and more information about A#B#C confounded with Columns, suppose the characters $B + C$ and $A + B + C$ are chosen. The unit character is A. The transpose of the 4×4 quasi-Latin square design in Table 1 is such a design, which is given in the first four columns of Design 2 in Table 8.

In combining the 4×2 rectangle and the 4×4 square, assign the values of the row characters in each row of the combined design so that they differ between the two segments. The last two columns of Design 2 in Table 8 have the quasi-Latin rectangle. The last column of the table applies to the last two columns of the design. The canonical efficiency factors for the combined design are those of Design 2 in Table 9 and it happens that they are the same as those for the design produced by CycDesigN 4.0. However, the two designs are not isomorphic.

Design 3: This design is constructed in the same manner as Design 1, but using different characters. To completely confound A#B#C with Columns, take $U = A + C$, $V = B + C$ and $W = X = Y = A + B + C$.

Table 8. Designs for a 2^3 factorial experiment in 4 rows $\times$ 6 columns

Design 1 — Method 1

$A+C$		$B+C$		$A+B+C$	
$=0$	$=1$	$=0$	$=1$	$=0$	$=1$
0, 0, 0	0, 0, 1	1, 0, 0	1, 0, 1	0, 1, 1	0, 1, 0
1, 0, 1	1, 0, 0	0, 1, 1	0, 1, 0	1, 1, 0	1, 1, 1
0, 1, 0	0, 1, 1	1, 1, 1	1, 1, 0	0, 0, 0	0, 0, 1
1, 1, 1	1, 1, 0	0, 0, 0	0, 0, 1	1, 0, 1	1, 0, 0

Design 2 — Method 3

	$B+C$		$A+B+C$		$A+B+C$			
	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$A+B^{\dagger}$	$A+C^{\dagger}$
$A+B=0$	1, 1, 1	1, 1, 0	0, 0, 0	0, 0, 1	0, 1, 1	1, 0, 0	$=1$,	$=1$
$A+B=1$	1, 0, 0	1, 0, 1	0, 1, 1	0, 1, 0	0, 0, 0	1, 1, 1	$=0$,	$=0$
$A+C=0$	0, 0, 0	0, 1, 0	1, 0, 1	1, 1, 1	1, 1, 0	0, 0, 1	$=0$,	$=1$
$A+C=1$	0, 1, 1	0, 0, 1	1, 1, 0	1, 0, 0	1, 0, 1	0, 1, 0	$=1$,	$=0$

†These relations apply only to units in the last two columns of the design

Table 9. Canonical efficiency factors and Residual degrees of freedom (DF) for the designs for a 2^3 factorial experiment in 4 rows × 6 columns

		Treatment sources							*Residual*
Design	*Unit sources*	A	B	C	A#B	A#C	B#C	A#B#C	*DF*
1	Rows	$\frac{1}{9}$	$\frac{1}{9}$	0	$\frac{1}{9}$	0	0	0	0
	Columns	0	0	0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	2
	Rows#Columns	$\frac{8}{9}$	$\frac{8}{9}$	1	$\frac{8}{9}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	8
2	Rows	0	0	0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	0	0
	Columns	0	0	0	0	0	$\frac{1}{3}$	$\frac{2}{3}$	3
	Rows#Columns	1	1	1	$\frac{8}{9}$	$\frac{8}{9}$	$\frac{5}{9}$	$\frac{1}{3}$	8
3	Rows	0	0	0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	0	0
	Columns	0	0	0	0	0	0	1	4
	Rows#Columns	1	1	1	$\frac{8}{9}$	$\frac{8}{9}$	$\frac{8}{9}$	0	9

The canonical efficiency factors and Residual degrees of freedom for the three designs are in Table 9. Design 3 has the advantage over the other designs in having more Residual degrees of freedom for Columns. To achieve this there is no information about A#B#C confounded with Rows#Columns. This is the design that was used in the experiment described by Tran (2009), but, in retrospect, Design 1 would have been better. The reason is that Design 1 has more information about the three-factor interaction confounded with Rows#Columns, and so is better able to distinguish between models with and without the three-factor interaction, with little loss of information about the other treatment effects from Rows#Columns.

An alternative to Design 1, when three-factor interactions are likely, is to ignore the factorial structure in the construction and to generate a design for eight treatments. For example, treatments 1–8 are assigned to the combinations of A, B and C listed in lexicographical order and CycDesigN 2.0 used to produce a design for the 8 treatments. The canonical efficiency factors for the factorial effects, when confounded with Rows#Columns, are more uniform than those for Designs 1–3, but the gain in efficiency for A#B#C is less than 10%. Also, the design does not have orthogonal factorial structure and cannot be made to have, no matter how the 2^3 factorial combinations are assigned to treatments 1–8.

Section 7.3 investigates constructing a row-column design when the number of replicates is increased from 3 to 5.

7.3. A 2^3 factorial in a 4×10 rectangle

For this example $p = 2$, $m = 3$, $k = 4$, $\ell = 10$ and $r = 5$. As always, Method 1 applies. Neither ℓ nor k are multiples of v and so Method 2 is not applicable. On the other hand, ℓ is neither a multiple of v nor a power of p so that Method 3 can be used. The design is an extended quasi-Latin rectangle design. We compare the three designs given in Table 10, that are in order of decreasing information about A#B#C confounded with Rows#Columns. They are constructed as follows:

Design 1: Method 3 is used and segments the design into a column segment of shape 4×8 and a second of shape 4×2 that contains an additional grid. The first segment uses Design 1 from Section 7.1 and the second segment is constructed using Method 1. It uses $A + B$ and $A + C$, and hence $B + C$, for row characters and $A + B + C$ for the column character.

Design 2: Either CycDesigN 2.0 or 4.0 (Whitaker et al., 2002; CycSoftware Ltd, 2009) is used, although the resulting design is not isomorphic to any of those constructed by our methods. A design with the same (partial) confounding of treatment effects as the computer-generated design can be constructed using Method 3 in the manner of Design 1. Here, the 4×8 segment is the same as Design 2 from Section 7.1, except that $B + C$ is used instead of $A + C$ in one of the column pairs. The 4×2 segment is the same as for Design 1.

Design 3: Method 1 is used, dividing the 4×10 rectangle into 1 row super-frame and 5 column super-frames. Each column super-frame contains a single column frame which is a grid of shape 4×2. The set of row characters is $\{A + B, A + C, B + C\}$; a 4×5 extended Latin square is used as an auxiliary design to assign the 4 pairs of treatments defined by the row characters. The column character is $A + B + C$ for all 5 column super-frames.

The canonical efficiency factors and Residual degrees of freedom for the three designs are given in Table 11. It appears that Designs 2 and 3 are suitable for situations in which it is appropriate to confound most of the three-factor interaction with Columns. In this circumstance, Design 3 has the advantage over Design 2 that there is no two-factor interaction confounded with Columns so that the rather small Residual degrees of freedom are increased by one. Design 1 would be preferred where the variance of the estimate of the three-factor interaction is to be minimized and the researcher is prepared to sacrifice some precision in estimating the two-factor interactions by partially confounding them with Columns; even so, only 24% of each two-factor interaction is confounded with Rows or Columns.

Table 10. Designs for a 2^3 factorial experiment in 4 rows × 10 columns

Design 1 — Method 3

$A+B$		$A+C$		$B+C$		$A+B+C$		$A+B+C$		
$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	*Relations*
0, 0, 0	1, 0, 0	0, 1, 0	0, 0, 1	0, 1, 1	1, 1, 0	1, 0, 1	1, 1, 1	0, 0, 0	1, 1, 1	$A+B=0,\ A+C=0^{\dagger}$
1, 1, 0	1, 0, 1	0, 0, 0	1, 0, 0	1, 1, 1	0, 0, 1	0, 1, 1	0, 1, 0	1, 0, 1	0, 1, 0	$A+B=1,\ A+C=0^{\dagger}$
0, 0, 1	0, 1, 0	1, 1, 1	0, 1, 1	0, 0, 0	1, 0, 1	1, 1, 0	1, 0, 0	1, 1, 0	0, 0, 1	$A+B=0,\ A+C=1^{\dagger}$
1, 1, 1	0, 1, 1	1, 0, 1	1, 1, 0	1, 0, 0	0, 1, 0	0, 0, 0	0, 0, 1	0, 1, 1	1, 0, 0	$A+B=1,\ A+C=1^{\dagger}$

Design 2 — Generated using CycDesigN: columns re-ordered to show confounding

$A+B+C$		$A+B+C$		$A+B+C$		$A+B+C$		$B+C$	
$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$
0, 1, 1	0, 1, 0	1, 0, 1	1, 0, 0	0, 1, 1	1, 1, 1	0, 0, 0	0, 0, 1	1, 0, 0	1, 1, 0
1, 1, 0	1, 1, 1	1, 1, 0	0, 1, 0	0, 0, 0	0, 0, 1	1, 0, 1	1, 0, 0	0, 1, 1	0, 0, 1
0, 0, 0	1, 0, 0	0, 0, 0	0, 0, 1	1, 1, 0	0, 1, 0	0, 1, 1	1, 1, 1	1, 1, 1	1, 0, 1
1, 0, 1	0, 0, 1	0, 1, 1	1, 1, 1	1, 0, 1	1, 0, 0	1, 1, 0	0, 1, 0	0, 0, 0	0, 1, 0

Design 3 — Method 1

$A+B+C$		$A+B+C$		$A+B+C$		$A+B+C$		$A+B+C$	
$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$
0, 0, 0	1, 1, 1	1, 1, 0	0, 0, 1	0, 1, 1	1, 0, 0	1, 0, 1	0, 1, 0	0, 0, 0	1, 1, 1
1, 0, 1	0, 1, 0	0, 0, 0	1, 1, 1	1, 1, 0	0, 0, 1	0, 1, 1	1, 0, 0	1, 0, 1	0, 1, 0
1, 1, 0	0, 0, 1	0, 1, 1	1, 0, 0	1, 0, 1	0, 1, 0	0, 0, 0	1, 1, 1	1, 1, 0	0, 0, 1
0, 1, 1	1, 0, 0	1, 0, 1	0, 1, 0	0, 0, 0	1, 1, 1	1, 1, 0	0, 0, 1	0, 1, 1	1, 0, 0

†These relations apply only to units in the last two columns of the design

Table 11. Canonical efficiency factors and Residual degrees of freedom (DF) for the designs for a 2^3 factorial experiment in 4 rows × 10 columns

		Treatment sources							*Residual*
Design	*Unit sources*	A	B	C	A#B	A#C	B#C	A#B#C	*DF*
1	Rows	0	0	0	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	0	0
	Columns	0	0	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{2}{5}$	5
	Rows#Columns	1	1	1	$\frac{19}{25}$	$\frac{19}{25}$	$\frac{19}{25}$	$\frac{3}{5}$	20
2	Rows	0	0	0	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	0	0
	Columns	0	0	0	0	0	$\frac{1}{5}$	$\frac{4}{5}$	7
	Rows#Columns	1	1	1	$\frac{24}{25}$	$\frac{24}{25}$	$\frac{19}{25}$	$\frac{1}{5}$	20
3	Rows	0	0	0	$\frac{1}{25}$	$\frac{1}{25}$	$\frac{1}{25}$	0	0
	Columns	0	0	0	0	0	0	1	8
	Rows#Columns	1	1	1	$\frac{24}{25}$	$\frac{24}{25}$	$\frac{24}{25}$	0	21

8. Choosing a unit structure

In the previous Sections we have constructed just row-column designs. However, as outlined in Section 3, no one type of design applies in all experimental situations and for quasi-Latin designs there is generally the option of using row-column, nested or contiguous designs for any specific experiment. That is, it is necessary, for any quasi-Latin design, to decide whether to employ a row-column design or an r-resolved design with the design divided into s whole frames and, if so, whether to latinize the divided design across its whole frames. This amounts to choosing between unit structures, which give different decompositions according to the unit sources. In addition, there is the question of how to construct r-resolved designs using our techniques. Choosing between different types of designs and the construction of r-resolved designs are illustrated for just two of the examples from Sections 5 and 7: the 2^3 factorial in an 4×8 square from Section 7.1 and the 2^5 factorial in an 8×8 square from Section 5.2.

This division of a design into several Latin squares or rectangles is fundamentally different from the division of a whole frame into frames or segments. Firstly, the division of a whole frame is purely a device for construction and is constrained by the construction algorithm. On the other hand, the division into several whole frames is done deliberately by the designer and is based on expected sources of variability in the experiment. Secondly, the division of whole frames is ignored in the randomization and analysis of the design. On the other hand, the formation of several whole frames results in randomizations between and, usually, within whole frames, depending on what is appropriate for the unit structure. It also results in the inclusion of terms corresponding to the whole frames in the analysis.

In constructing both nested and contiguous designs the first step is to specify the number and size of whole frames, which is akin to Method 3, except that the designer has more freedom in choosing the size of the whole frames. However, each whole frame needs to meet the conditions for a quasi-Latin square or (extended) quasi-Latin rectangle. For nested designs, the second step is to apply the methods we have presented to each whole frame independently, although the overall pattern of (partial) confounding of the treatment effects must be considered. For contiguous designs, the whole frames are joined into a single frame and the methods that we have presented are applied to this combined frame. This ensures that the same characters are confounded between the contiguous entities (rows or columns). However, in choosing the characters for the noncontiguous entities, the overall pattern of their (partial) confounding must be considered. Also, care is needed in choosing the unit characters and auxiliary design Δ_3, as this will determine the treatment effects confounded with the interaction of whole frames and the contiguous entities.

A design constructed as one type can often be deployed as a design of a different type. This requires the randomization and analysis appropriate to the type of design actually deployed. For example, a Latin square design can be deployed as a randomized complete block design, by randomizing and analysing the Latin square design for one blocking factor nested within the other. That is, the unit structure differs between the constructed and deployed designs. Thus, the nested designs constructed in this Section can be deployed as row-column designs and the contiguous designs as row-column or nested designs.

8.1. *The* 2^3 *factorial in a* 4×8 *rectangle revisited*

In Section 5.1 it was suggested that two 4×4 squares, like the plan given by Cochran and Cox (1957, Table 8.1), are more useful than a single square. It was also noted that constructing a single 4×8 rectangle, as is done in Section 7.1, is an alternative. Here, this alternative and two nested and two contiguous quasi-Latin square designs with two squares are compared. For the designs constructed here, the first step is to divide them into two squares (whole frames) of shape 4×4. Then, $p = 2$, $m = 3$, $s = 2$, $k = \ell = 4$ and $r = 2$ so $t = u = 2$ and $c = d = 2$.

The two nested designs are constructed by applying Method 1 to each square. Design 1 is the quasi-Latin square design given in Cochran and Cox (1957, Table 8.1). The first square of the design is in Table 1; the second square is obtained from this by swapping row and column characters. The design involves complete replicates in grids of shape 2×4 and 4×2 in each square, and is a nested, 2-resolved design.

Design 2, also a nested, 2-resolved design, is constructed in the same way as Design 1. For its first square, $U = V = A + B + C$, $W = A + B$, $X = A + C$ and $Y = A$; the second square uses the same characters, except that $X = B + C$ and $Y = B$.

These nested designs do not take advantage of the contiguity of the rows of the two squares because there is no constraint on the treatments assigned to the same row in different squares. On the other hand, although the quasi-Latin rectangle design in Table 6 has a complete replicate in each row and in each of four 4×2 grids, it also not suitable as a contiguous design, because no attention has been paid to the confounding with rows within squares. In particular, it does not have orthogonal factorial structure. Our construction method can be used to choose a better confounding pattern for a contiguous design.

Design 3 consists of two row-contiguous 4×4 quasi-Latin squares and is 2-resolved. To construct it, we apply Method 1 to the whole design, which is of shape 4×8. The construction is similar to that of Design 3 in Section 7.1. That is, we require four column characters, which need not be different, and one set of three unit characters that are closed under addition. Also, necessary is an auxiliary design Δ_3 for assigning the values of the unit characters. For example, take as the column characters $B+C$ and $A+C$ in both squares to leave other interaction characters for unit characters. Take the set $\{A + B + C, A + B, C\}$ for unit characters. In order to have $A+B+C$ and $A+B$, but not C, partially confounded with the Rows#Squares, number the combinations of the values of the first two characters as follows: $1 = (0, 0)$, $2 = (0, 1)$, $3 = (1, 0)$ and $4 = (1, 1)$. Then assign assign these groups to the 4×4 array of 1×2 subframes using the particular Latin square whose rows are $(2, 1, 3, 4)$, $(3, 4, 2, 1)$, $(1, 3, 4, 2)$ and $(4, 2, 1, 3)$. The design is in Table 12.

CycDesigN cannot produce a design with the same unit structure as Design 3, because it generates only 1-resolved designs. Instead, Design 4 is a contiguous design for a 2×2 array of 2×4 grids produced using CycDesigN; the design is resolved and latinization is across both rows and columns. The unit factors are BigRows, BigCols, Rows and Columns, all with 2 levels except that Columns has 4 levels. Also, they all are crossed, except that Rows is nested within BigRows and Columns is nested BigCols. The design is in Table 12.

Table 12. Contiguous designs for a 2^3 factorial experiment in 4 rows $\times$ 8 columns

Design 3 — Method 1							
$B+C$		$A+C$		$B+C$		$A+C$	
$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$
0, 1, 1	1, 0, 1	0, 0, 0	1, 1, 0	1, 1, 1	0, 0, 1	0, 1, 0	1, 0, 0
1, 1, 1	0, 0, 1	0, 1, 0	1, 0, 0	0, 1, 1	1, 0, 1	0, 0, 0	1, 1, 0
0, 0, 0	1, 1, 0	1, 1, 1	0, 0, 1	1, 0, 0	0, 1, 0	1, 0, 1	0, 1, 1
1, 0, 0	0, 1, 0	1, 0, 1	0, 1, 1	0, 0, 0	1, 1, 0	1, 1, 1	0, 0, 1
Design 4 — Generated using CycDesigN							
0, 1, 0	0, 1, 1	0, 0, 1	0, 0, 0	1, 1, 1	1, 0, 0	1, 1, 0	1, 0, 1
1, 1, 1	1, 1, 0	1, 0, 0	1, 0, 1	0, 0, 0	0, 1, 1	0, 0, 1	0, 1, 0
1, 1, 0	0, 1, 0	0, 1, 1	1, 1, 1	1, 0, 1	0, 0, 0	1, 0, 0	0, 0, 1
1, 0, 1	0, 0, 1	0, 0, 0	1, 0, 0	0, 1, 1	1, 1, 0	0, 1, 0	1, 1, 1

The canonical efficiency factors and Residual degrees of freedom for the four designs are given in Table 13. As described by Cochran and Cox (1957), Design 1 has a quarter of the information of the effects for each of $\{A+B, A+C, B+C, A+B+C\}$ confounded with both rows and columns. Main effects are orthogonal to rows and columns. Design 2, compared with Design 1, has A#B#C completely confounded with Rows [Squares] and more information about the two-factor interactions confounded with Rows # Columns [Squares]. However, the effect of removing Squares variability would appear to be a decrease in the amount of information about the interactions in the lowest unit source; both designs have lower efficiencies for the interactions than those for Design 1 in Section 7.1. Of the contiguous designs, Design 3 is better than Design 4, because it has orthogonal factorial structure, and, for the last unit source, it has (i) main effects completely confounded with it, (ii) more three-factor information partially confounded with it, and (iii) more Residual degrees of freedom.

Designs 3 and 4, while constructed as contiguous designs, could also be deployed as row-column or nested designs. The analysis for Design 3 when deployed as a nested de-

Table 13. Canonical efficiency factors and Residual degrees of freedom (DF) for the nested and contiguous designs for a 2^3 factorial experiment in 4 rows $\times$ 8 columns

		Treatment sources							*Residual*
Design	*Unit sources*	A	B	C	A#B	A#C	B#C	A#B#C	*DF*
1	Squares	0	0	0	0	0	0	0	1
	Rows [Squares]	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	2
	Columns [Squares]	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	2
	Rows # Columns [Squares]	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	11
2	Squares	0	0	0	0	0	0	0	1
	Rows [Squares]	0	0	0	0	0	0	1	5
	Columns [Squares]	0	0	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	3
	Rows # Columns [Squares]	1	1	1	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{3}{4}$	0	12
3	Squares	0	0	0	0	0	0	0	1
	Rows	0	0	0	0	0	0	0	3
	Rows # Squares	0	0	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	1
	Columns [Squares]	0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0	4
	Rows # Columns [Squares]	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	11
4	BigRows	0	0	0	0	0	0	0	1
	BigCols	0	0	0	0	0	0	0	1
	BigRows#BigCols	0	0	0	0	0	0	0	1
	Rows [BigRows]	0	0	0	0	0	0	0	2
	Columns [BigCols]	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	0	$0^\dagger$	$\frac{1}{8}^\dagger$	$0^\dagger$	2
	Rows # BigCols [BigRows]	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0	0	0
	Columns # BigRows [BigCols]	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$0^\dagger$	$\frac{1}{8}^\dagger$	$\frac{1}{2}^\dagger$	0
	Rows # Columns [BigRows $\wedge$ BigCols]	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	5

†This treatment source is nonorthogonal to previous treatment sources estimated from the same unit source and its canonical efficiency factor is adjusted for the previous sources.

sign combines the Rows and Rows#Squares unit sources from the analysis of the contiguous design. The nested design has the advantage over the contiguous design of having more Residual degrees of freedom for Rows, but would subject A#C and A#B#C to Rows variability. Hence, Design 3 should not be deployed as a nested design when appreciable differences between long rows are anticipated. The efficiencies for the last unit source for Design 3, as a nested design, and for Design 1 are equal. However, of these two nested designs, Design 3 has the advantage that it has more Residual degrees of freedom for Rows [Squares] and Columns [Squares]. The analysis for Design 3 when deployed as a row-column design can be obtained from its analysis as a contiguous design by combining (i) the Squares and Columns [Squares] sources and (ii) the Rows#Squares and Rows # Columns [Squares] sources. The result is that A#B and A#B#C are now completely confounded with Rows#Columns, the latter having 14 Residual degrees of freedom. Using Design 3 as a row-column design is restricted to situations in which it is anticipated that squares are similar and the Rows#Squares is not appreciable source of variability.

Design 1 in Section 7.1 and Designs 1 and 3 from this Section are row-column, nested and contiguous designs, respectively, for which the amount of information about A#B#C partially confounded with the last unit source is maximized for each type of design. Comparing them shows that allowing for the removal of the difference between Squares reduces the last unit source's (i) efficiencies for the interactions and (ii) Residual degrees of freedom (from 14 to 11). As suggested in Section 3, the choice between these designs depends on the sources of unit variability that are expected. One might argue that there is little to be lost in routinely employing a contiguous design because, if unit sources are shown to be negligible in a preliminary analysis, then one can drop these unit sources and so increase the efficiencies and Residual degrees of freedom associated for the remaining sources in the final analysis. However, as noted in Section 3, this is a strategy with undesirable consequences and it is preferable to use the analysis appropriate to the design chosen. Further, the contiguous design, when analysed as a row-column design, is inferior to Design 1 from Section 7.1.

8.2. The 2^5 factorial in an 8×8 square revisited

Section 5.2 describes a quasi-Latin square design for arranging a 2^5 factorial in an 8×8 square. The design in Section 5.2, as constructed, is resolvable; grids, each containing a complete set of treatments, are obtained by dividing either the rows into two 4×8 grids or the columns into two 8×4 grids. So again, while the design was formulated as a row-column design, it could be deployed as either a nested or a contiguous design. The resulting design is a quasi-Latin rectangle design. However, as in Section 8.1, the characters used in constructing the particular design given in Section 5.2 are not ideal for these other unit structures and so we again employ our construction method to choose a better confounding pattern.

The first step in constructing either a nested or a contiguous design is to divide the design into whole frames. We consider a design in which there are two rectangles (whole frames) of shape 4×8, each of which is a grid. Thus, $p = 2$, $m = 5$, $s = 2$, $k = 4$, $\ell = 8$ and $r = 1$.

Constructing a nested quasi-Latin rectangle design for two such rectangles is a straightforward application of Method 1 to each rectangle. Different column characters can be chosen in the two rectangles.

We now give two column-contiguous quasi-Latin rectangle designs consisting of two rectangles of shape 4×8. Design 1 is constructed by applying Method 1 to the whole design. So, as for the design in Section 5.2, the new design consists of a 2×2 array of 4×4 sub-squares, and two sets of row characters and two sets of columns characters are required. Each set contains three characters closed under addition (modulo 2). Also, a unit character is required and this must be chosen carefully as it, and its sums with the column characters, are confounded with Columns#Grids. The chosen sets of row characters are $\{A+B+C,\ C+D+E,\ A+B+D+E\}$ and $\{A+B+C+E,\ B+C+D+E,\ A+D\}$ and the sets of column characters are $\{A+B+C+D,\ A+C+E,\ B+D+E\}$ and $\{A+C+D+E,\ B+C+D,\ A+B+E\}$. The unit character is $B+C+E$. It is selected because none of its sums with column characters results in a main effect.

Design 2 was constructed using CycDesigN 4.0 and it is also a column-contiguous design. It was obtained using default weights in a two-stage search in which each stage was allowed to run for between 210 and 420 seconds on a computer running Windows XP. If one stops the searches sooner, as discussed in Section 7.1, a design with lower weighted efficiency will be obtained.

The skeleton analysis-of-variance tables for Designs 1 and 2 are in Table 14. They are divided into subtables according to the unit sources. The treatment sources confounded with a unit source, along with their efficiencies and degrees of freedom, are incorporated into the subtable for that unit source. An important difference between the two designs is that Design 1 has orthogonal factorial structure whereas Design 2 does not. As a result, the estimation of one treatment effect is independent of another for Design 1, but not for Design 2. Design 1 also has an advantage over Design 2 in the estimation of two-factor interactions. Unlike Design 2, Design 1 has no two-factor interactions partially confounded with Columns; it confounds more with Columns#Grids, which is expected to be less variable than Columns. The two designs have a similar amount of information about two-factor interactions confounded with Rows # Columns [Grids]. Design 1 has more information about three-factor interactions confounded with Rows # Columns [Grids].

Design 3 is constructed like Design 1, but using different characters, so that all main effects are estimated with full efficiency from the last unit source and all two- and three-factor interactions from the last unit source with as high efficiency as possible. The sets of row characters are $\{A+C+D,\ B+C+E,\ A+B+D+E\}$ and $\{B+C+D,\ A+C+E,\ A+B+D+E\}$ and the sets of column characters are $\{A+B,\ A+D+E,\ B+D+E\}$ and $\{D+E,\ A+B+D,\ A+B+E\}$. The unit character is $A+B+C+D+E$.

Like Designs 3 and 4 in Section 8.1, these contiguous designs can be deployed as row-column, nested or contiguous designs, with similar considerations to those outlined in Section 8.1.

Table 14. Skeleton analysis-of-variance tables for a 2^5 factorial experiment in 2 column-contiguous rectangles of 4 rows × 8 columns (G = Grids; R = Rows; L = Columns)

units		*treatments*					
		Design 1			*Design 2*		
Source	*DF*	*Eff.*§	*Source*	*DF*	*Eff.*§	*Source*¶	*DF*
Mean	1	1	Mean	1	1	Mean	1
Grids	1						
Rows [G]	6	$\frac{1}{2}$	A#D	1	$\frac{1}{8}$	B#D	1
		$\frac{1}{2}$	A#B#C	1	$\frac{1}{2}$	A#B#C	1
		$\frac{1}{2}$	C#D#E	1	$\frac{1}{8}$	A#B#D	1
		$\frac{1}{2}$	A#B#C#E	1	$\frac{1}{8}$	A#B#E	1
		$\frac{1}{2}$	A#B#D#E	1	$\frac{1}{4}$	A#D#E†	1
		$\frac{1}{2}$	B#C#D#E	1	$\frac{1}{4}$	C#D#E	1
Columns	7	$\frac{1}{2}$	A#B#E	1	$\frac{1}{4}$	A#C	1
		$\frac{1}{2}$	A#C#E	1	$\frac{1}{4}$	B#C	1
		$\frac{1}{2}$	B#C#D	1	$\frac{1}{8}$	A#D	1
		$\frac{1}{2}$	B#D#E	1	$\frac{1}{16}$	A#E†	1
		$\frac{1}{2}$	A#B#C#D	1	$\frac{1}{8}$	D#E†	1
		$\frac{1}{2}$	A#C#D#E	1	$\frac{1}{8}$	A#B#D	1
			Residual	1	$\frac{1}{8}$	A#C#D	1
L # G	7	$\frac{1}{2}$	A#B	1	$\frac{1}{4}$	A#C	1
		$\frac{1}{2}$	A#C	1	$\frac{1}{4}$	B#C	1
		$\frac{1}{2}$	C#D	1	$\frac{1}{8}$	A#D	1
		$\frac{1}{2}$	D#E	1	$\frac{1}{16}$	A#E†	1
		$\frac{1}{2}$	A#B#D	1	$\frac{1}{8}$	D#E†	1
		$\frac{1}{2}$	A#D#E	1	$\frac{5}{8}$	A#B#D	1
		1	B#C#E	1	$\frac{1}{40}$	A#C#D†	1
R # L [G]	42	1	Main effects	5	1	Main effects	5
		0.67	Two-factor	10	0.68	Two-factor‡	10
		0.53	Three-factor	9	0.21	Three-factor‡	10
		0.50	Four-factor	5	0.56	Four-factor‡	5
		1	A#B#C#D#E	1	$\frac{1}{2}$	A#B#C#D#E†	1
			Residual	12		Residual	11

§For single-degree-of-freedom sources, the efficiencies are the canonical efficiency factors; when a source is nonorthogonal to previous treatment sources estimated from the same unit source, its canonical efficiency factor is adjusted for the previous sources. Those for the sources with multiple degrees of freedom are the harmonic means of the canonical efficiency factors or the A-optimality criterion (John and Williams, 1995, Section 2.4).

†This source is nonorthogonal to previous treatment sources estimated from the same unit source.

‡Not all these interactions are orthogonal to each other.

¶The following sources are partially confounded with the accompanying unit source, but there is no information about them remaining after the previous treatment sources have been fitted:

Rows [G]**:** B#E, C#D, C#E, A#C#D, A#C#E, A#B#D#E, B#C#D#E and A#B#C#D#E;

Columns**:** A#B#E, A#C#E, B#C#D, B#C#E, B#D#E, C#D#E, A#B#C#D, A#B#C#E and A#C#D#E;

L # G**:** A#B#E, A#C#E, B#C#D, B#C#E, B#D#E, C#D#E, A#B#C#D, A#B#C#E and A#C#D#E.

9. Discussion

The (extended) quasi-Latin rectangle designs increase the range of situations in which a row-column design with orthogonal factorial structure can be employed for assigning factorial treatments. In particular, they allow for more choice in the number of replicates of each treatment than is available with quasi-Latin square designs.

The construction methods that we present are flexible and permit a degree of direct control of the confounding in a quasi-Latin design, as shown in Section 7. For example, in Sections 7.1 and 7.2, all designs can be produced using our construction methods and the designer can choose which two-factor interactions are partially confounded with Columns. It would be helpful to practitioners if software assisted in this by searching through possible confounding characters and listing them, in the way that PLANOR (Kobilinsky, 1994) currently does for fractional factorial designs.

The methods can be used to produce row-column, nested or contiguous designs as demonstrated in Section 8. A contiguous design can always be deployed as a row-column or a nested design, by varying the randomization and analysis to suit the deployed design. However, the deployed design may not be the best of its type.

The practice of habitually employing a contiguous design on the basis that nonsignificant unit sources can be dropped after a preliminary analysis is discouraged on the grounds that this would mean that (i) often the best design is not used and (ii) an analysis strategy with undesirable consequences is employed. Our recommendation is that, in designing an experiment, the designer identify the expected sources of unit variability for the experiment and these determine the unit structure. Then a purpose-built design is constructed of the type corresponding to the unit structure for the experiment. The chosen design should be randomized, and the results analysed, according to its unit structure.

Generating quasi-Latin designs using computer search algorithms, such as the one employed by CycDesigN, may not produce a design with the properties desired by the designer. For example, with CycDesigN as currently implemented, the control over the confounding of two-factor interactions is not as flexible as with our construction method. Also, the design produced can depend upon the seed for the random number generator and the length of the search time. There is also no guarantee that computer-generated designs will have orthogonal factorial structure. On the other hand, software like CycDesigN has the distinct advantages that it can produce designs for a wider range of design parameters than our construction methods and that it is easy to use. Also, a design produced by CycDesigN, if one has access to it, can be used as a benchmark for constructed designs, as was done by Tran (2009).

Acknowledgments

We are very grateful to Henry Mancinni and Joan Gibbs for providing the motivating experimental situation that led to this work. We would also like to acknowledge the financial support of various institutions that supported parts of this work: T. T. Tran received a studentship from Co-operative Research Centre of Railway Engineering and Technologies (Rail CRC), Australia; R. A. Bailey undertook work while on a Sir Frederick McMaster Fellowship funded by CSIRO, Australia; C. J. Brien and R. A. Bailey received funding from EPSRC to attend the 2011 Programme on Design and Analysis of Experiments at the Isaac Newton Institute for the Mathematical Sciences where further work was carried out.

References

Bailey, R. A. (1985). Balance, orthogonality and efficiency factors in factorial design. *Journal of the Royal Statistical Society, Series B (Methodological) 47*, 453–458.

Bailey, R. A. (2008). *Design of Comparative Experiments*. Cambridge: Cambridge University Press.

Brien, C. J. and R. A. Bailey (2009). Decomposition tables for experiments. I. A chain of randomizations. *The Annals of Statistics 37*, 4184–4213.

Brien, C. J., B. D. Harch, R. L. Correll, and R. A. Bailey (2011). Multiphase experiments with at least one later laboratory phase. I. Orthogonal designs. *Journal of Agricultural, Biological and Environmental Statistics 16*, 422–450.

Cochran, W. G. and G. M. Cox (1957). *Experimental Designs* (2nd ed.). New York: John Wiley & Sons.

CycSoftware Ltd (2009). CycDesigN 4.0: A package for the computer generation of experimental designs. CycSoftware Ltd, New Zealand.

Edmondson, R. N. (1989). Glasshouse design for repeatedly harvested crops. *Biometrics 45*, 301–307.

Gilmour, S. G. and P. Goos (2009). Analysis of data from non-orthogonal multistratum designs in industrial experiments. *Journal of the Royal Statistical Society, Series C (Applied Statistics) 58*, 467–484.

Healy, M. J. R. (1951). Latin rectangle designs for 2^n factorial experiments on 32 plots. *The Journal of Agricultural Science 41*, 315–316.

Janky, D. G. (2000). Sometimes pooling for analysis of variance hypothesis tests: A review and study of a split-plot model. *The American Statistician 54*, 269–279.

John, J. A. and E. R. Williams (1995). *Cyclic and Computer Generated Designs* (2nd ed.). London: Chapman and Hall.

Kobilinsky, A. (1994). PLANOR: programme de génération automatique de plans d'experiénces réguliers. Unité de Biométrie (BIA), INRA, F78352. Jouy-enJosas Cedex, France.

Preece, D. A. (2006). Latin squares, latin cubes, latin rectangles. In S. Kotz, C. B. Read, N. Balakrishnan, and B. Vidakovic (Eds.), *Encyclopedia of Statistical Sciences*. New York: John Wiley & Sons. DOI: <http://dx.doi.org/10.1002/0471667196.ess1400.pub2>.

Rao, C. R. (1946). Confounded factorial designs in quasi-Latin squares. *Sankhyā 7*, 295–304.

Shah, S. M. (1978). On α-resolvability and affine α-resolvability of incomplete block designs. *The Annals of Statistics 6*, 468–471.

Tran, T. T. (2009). *Design and Analysis of Experiments for Assessing Indigenous Plant Species*. Ph. D. thesis, School of Mathematics and Statistics, The University of South Australia, Adelaide, South Australia.

Whitaker, D., E. R. Williams, and J. A. John (2002). CycDesigN: A package for the computer generation of experimental designs (Version 2.0). CSIRO, Canberra, Australia.

Williams, E. R. (1986). Row and column designs with contiguous replicates. *Australian Journal of Statistics 28*, 154–163.

Williams, E. R. and J. A. John (1996). Row-column factorial designs for use in agricultural field trials. *Journal of the Royal Statistical Society, Series C (Applied Statistics) 45*, 39–46.

Williams, E. R., A. C. Matheson, and C. E. Harwood (2002). *Experimental Design and Analysis for Tree Improvement* (2nd ed.). Melbourne, Australia: CSIRO.

Yates, F. (1937). The design and analysis of factorial experiments. *Imperial Bureau of Soil Science Technical Communication 35*.

Youden, W. J. (1940). Experimental designs to increase accuracy of greenhouse studies. *Contributions from Boyce Thompson Institute 11*, 219–228.