Exposing the confounding in experimental designs to understand and evaluate them, and formulating linear mixed models for analyzing the data from a designed experiment

Christopher James Brien*[,1,2], Renata Alcarde Sermarini[3] and Clarice Garcia Borges Demétrio[3]
[1]School of Agriculture, Food and Wine, University of Adelaide, PMB 1, Glen Osmond. SA 5064
[2]UniSA STEM, University of South Australia, GPO Box 2471, Adelaide. SA 5001
[3]Luiz de Queiroz College of Agriculture (ESALQ), University of São Paulo, Caixa Postal 9, Piracicaba

January 18, 2023

Abstract

Comparative experiments involve the allocation of treatments to units, ideally by randomization. This necessarily confounds treatment information with unit information, which we distinguish from the other forms of information blending, in particular aliasing and marginality. We outline a factor-allocation paradigm for describing experimental designs with the aim of (i) exhibiting the confounding in a design, using analysis-of-variance-like tables, so as to understand and evaluate the design and (ii) formulating a linear mixed model based on the factor allocation that the design involves. The approach exhibits the dispersal of treatments information between units sources, allows designers a choice in the strategy that they adopt for including block-treatment interactions, clarifies differences between experiments, accommodates systematic allocation of factors and provides a consolidated analysis of nonorthogonal designs. It provides insights into the process of designing experiments and issues that commonly arise with designs. The paradigm has pedagogical advantages and is implemented using the `R` package `dae`.

Keywords: block-treatment interaction; confounding; factor allocation; linear mixed model; repeated-measurements experiments
Supporting Information for this article is available from the author or on the WWW under `http://dx.doi.org/10.1022/bimj.XXXXXXX`

1 Introduction

An experimental design for a comparative experiment usually involves the random allocation of a set of treatments to a set of units in a manner that reduces the effects of large sources of unit variation on the estimation of the treatment effects. This can be formally cast as the selection of design that minimizes an optimality criterion under a model that has terms for the treatments and the sources of unit variation. The traditional method of determining the model is to list all the treatment and unit terms thought likely to be important in the experiment and to designate each term as being either fixed or random. A number of authors have stressed the importance of identifying treatment and either block or unit factors (Wilk and Kempthorne, 1956; Nelder, 1965a,b; Brien, 1989; Milliken and Johnson, 1992; Piepho et al., 2003; Littell et al., 2006; Stroup, 2012). However, the details differ between their implementations. What is generally not specifically addressed by these authors is the confounding that results from the allocation of treatments to units.

Here, a distinction is made between confounding and aliasing that is consistent with the usage

*Corresponding author: e-mail: chris.brien@adelaide.edu.au

originally established by Fisher (1935b) and Finney (1943) for factorial experiments and continues to be employed in modern experimental design textbooks (see for example Montgomery, 2013). While confounding was originally confined to the relationship of treatment contrasts to block contrasts, a more general view is taken here: *confounding* is the blending together of treatment contrasts and unit contrasts. A consequence of this more general view is that all experiments, whether randomized or systematic, involve confounding by virtue of the allocations they employ. On the other hand, *aliasing* is the mixing up of different treatment contrasts, or, less frequently, of different units contrasts. It was originally introduced in the context of fractional factorial experiments. In such experiments various factorial effects associated with the treatments are aliased with each other. It occurs in factorial experiments in which the replication of the combinations of the levels of the treatments factors are not (proportionally) balanced.

This paper describes an approach for designing, evaluating and model formulation for comparative experiments, called the *factor-allocation paradigm*, an objective of which is to identify the confounding in an experimental design. To do this, a design is analyzed to produce an anatomy table, that is like a skeleton analysis-of-variance (ANOVA) table but that gives a concrete depiction of the design's confounding. The approach has been developed by Brien and his co-authors for multiphase experiments (Brien et al., 2011; Brien, 2017, 2022). However, its applicability to the more usual, single-phase experiments is obscured in these publications by the complexities of multiphase experiments. Here, a simplified version is described and its practical benefits are stressed. The paradigm is based on the identification of the factors that are *allocated* (treatments) and those that are *recipient* (units) factors, this characterization of the factors emphasizing their roles in the assignment of treatments for an experimental design. As Brien (1983) suggests, this approach has its origins in the separation of the 'topographical' and 'treatment' structures identified by Fisher (1935a).

Section 2 uses simple experimental designs on a 5×5 grid of plots to introduce the factor-allocation paradigm and to demonstrate the role that the anticipated effects play in the design of an experiment. Section 3 considers the incorporation of block-treatment interactions in designing an experiment and contrasts the paradigm with traditional approaches. Using examples in which time is a factor, Section 4 illustrates several further insights that accrue from considering the factor allocation in a design. Section 5 demonstrates how the approach handles nonorthogonal designs. Section 6 discusses the practical benefits of the paradigm. Supporting information is available that provides (i) a glossary of the terminology used in the paper, (ii) figures showing the layouts and the `R` (R Core Team, 2023) scripts for the examples, and (iii) the theory on which the approach is founded.

2 Designing experiments on a 5×5 grid of plots

The process underlying the factor-allocation paradigm is depicted in Figure 1, which is a modified version of Brien (2022, Figure 1). The process begins with determining the *anticipated model* as a means to identifying the effects that are to be taken into account in the experimental design, preferably in concert with the researchers. This step formalizes the deliberations traditionally undertaken in designing an experiment. It can be facilitated by considering all of the factors in the experiment and identifying those that are the unit factors and those that are the treatment factors. The unit factors belong to the units and are, potentially at least, *recipient factors* in that they are to receive factors that are allocated to them. Generally, the treatment factors are the *allocated factors*, bearing in mind that not all *factors of interest* are allocated. In considering the terms for effects to be taken into account in a design, it is useful to identify the inherent crossing and nesting relationships between the recipient factors. Then it can be decided whether the factor allocation in the design being developed is to respect any crossing of unit factors or that, because some effects can be ruled out as being unlikely in the experiment, some crossed factors are to be treated as nested. The initial and homogeneous allocation models are called allocation models because they are based on terms appropriate to the allocations represented in the factor allocation diagram. Some common types of allocations are those that use randomization, restricted randomization, spatially optimal allocation or systematic allocation. When all allocations employ (restricted) randomization, the initial allocation model will be equivalent to a randomization model and randomization-based inference is available. The homogeneous allocation model is derived from the initial allocation model by changing the fixed/random designation of terms and by including block-treatment interactions. Thus, randomization inference is no longer possible. However, the models

retain all terms appropriate to the allocations performed in generating the layout for an experiment.

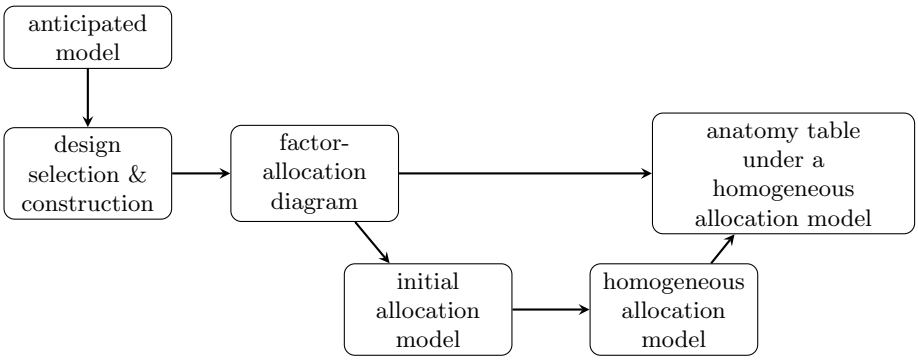

FIGURE 1 Flowchart of items produced in using the factor-allocation paradigm for the design of an experiment.

To illustrate the factor-allocation approach to designing experiments, consider an experiment in which five varieties are to be allocated to a 5×5 grid of 25 plots. The factors are, say, Rows, Columns and Varieties. For a grid of plots, Rows and Columns are intrinsically crossed. However, in this situation, one might employ a completely randomized design (CRD), a randomized complete-block design (RCBD) or a Latin square design (LSqD). Which of these three designs is appropriate depends on which, if any, of the Rows and Columns effects are expected to be present. The situation in which both effects cannot be ruled out is investigated first and then we examine the case when Columns effects are not expected.

2.1 Both Rows and Columns effects expected

The anticipated model. The paradigm begins with the specification of the anticipated model, which serves to capture formally the terms that need to be accounted for in selecting a design. The *symbolic anticipated model*, that uses *abbreviated term names* consisting of just the initial capital letters of the factor names for the terms, is, for this experiment when both Rows and Columns effects are expected:

$$\text{Mean} + \text{V} \quad | \quad \text{Mean} + \text{R} + \text{C} + \underline{\text{R:C}}, \tag{1}$$

where terms before the '|' are fixed and so form the *expectation model*, while those after it are random and form the *variance model*; the underlined term, called an *identity term*, uniquely indexes the observational units, the plots; the ':' signifies that this term is based on all observed combinations of the levels of the factors R(ows) and C(olumns). That both Rows and Columns main effects are in (1) means that it is inherent to this model that Rows and Columns are crossed. In this case, the crossing inherent to the model matches the intrinsic crossing of the physical layout of the units in a grid.

Each of the random terms is assumed to contribute a variance component, $\sigma^2_{\cdot}$, to the variance of the response. The fixed terms will contribute to the expectation of the response.

A term for the trivial factor (overall) Mean is included in both the expectation and variation models. It is needed in both models to allow for an expectation model with no treatment effects and to give a variance model whose eigenspaces are orthogonal. The Mean term in the expectation model represents the overall population mean for comparable experiments; the Mean term in the variance model represents the basic covariance between observations within experiments, or equivalently reflects the excess variation in the overall mean between comparable experiments over and above that within experiments.

A *formal marginal linear mixed model* for the experiment is:

$$\mathbb{E}[\mathbf{Y}] = \mu\mathbf{1} + \mathbf{X}_{\text{V}}\boldsymbol{\tau}_{\text{V}}, \text{var}[\mathbf{Y}] = \sigma^2_{\mu}\mathbf{S}_{\text{Mean}} + \sigma^2_{\text{R}}\mathbf{S}_{\text{R}} + \sigma^2_{\text{C}}\mathbf{S}_{\text{C}} + \sigma^2_{\text{RC}}\mathbf{S}_{\text{RC}}, \tag{2}$$

where $\mathbf{Y}$ is the vector for a random response variable, $\mathbf{1}$ is a vector of ones, $\mathbf{X}_{\text{V}}$ is an indicator matrix for the fixed Varieties term, $\mathbf{S}_{\text{Mean}}$, $\mathbf{S}_{\text{R}}$, $\mathbf{S}_{\text{C}}$ and $\mathbf{S}_{\text{RC}}$ are relationship matrices for the random terms ($\mathbf{S}_{\cdot} = \mathbf{Z}_{\cdot}\mathbf{Z}_{\cdot}^{\top}$, where $\mathbf{Z}_{\cdot}$ is the indicator matrix for the random term indicated by the subscript); μ and $\boldsymbol{\tau}_{\text{V}}$ are the fixed parameters for the Mean and for Varieties, and σ^2_{μ}, σ^2_{R}, σ^2_{C} and σ^2_{RC} are the variance components for the terms Mean, Rows, Columns and Rows:Columns. For the rest of the paper only symbolic models will be presented, these models being readily incorporated into scripts for model fitting.

Design selection & construction. Informally, a design that honours the crossing between Rows and Columns is required and an LSqD is such a design. Formally, a design that minimizes the average variance of the pairwise variety differences under the anticipated model is sought. For the examples in this paper, this implies that they are A-optimal, a property that is further elaborated upon in the Anatomy table section below. The LSqD is an orthogonal design under the anticipated model in (1) and is the A-optimal design under this model, or any alternative designation of Rows and Columns as either fixed or random (Shah and Sinha, 1989).

In general, a design for an experiment can be chosen using: (i) standard designs (e.g. CRD, RCBD, LSqD), (ii) a catalogue of designs (e.g. Cochran and Cox, 1957) or, (iii) computer searches. Software available for computer searching includes `CycDesigN` (VSN International, 2022a), `odw` (Butler, 2022), `JMP` (SAS Institute Inc., 2022) and `SAS PROC OPTEX` (SAS Institute Inc., 2020). However, to quote Mead et al. (2012, p.558), "It is our belief that ... packages and catalogues provide a relatively small contribution to the overall designing of practical experiments". On the other hand, the formal process of specifying an anticipated model and seeking a design that is A-optimal under it can be, and often is, dispensed with in practice, especially for approaches (i) and (ii); it could be for all of the examples in this paper. Nonetheless, it does make the underpinnings of design selection explicit and, when computer searches are employed, the anticipated model is needed, it being the model that is available.

A design is usually a systematic arrangement of the treatments to the units. A general method of randomizing a *systematic design* to obtain a *randomized layout* is to appropriately permute its units factors followed by a joint re-ordering of its treatments and units factors such that its units factors are returned to standard order. This method of randomization has been implemented in the `designRandomize` function from the `R` package `dae` (Brien, 2023) and is used for all the designs in this paper. The appropriate permutations for an LSqD are to permute (i) the Rows and (ii) the Columns. The `R` script in Supplement Section C.1 produces a randomized layout and plots it. The permutation chosen for Rows was (5,1,2,3,4) and for Columns was (3,1,2,4,5). These permutations are combined to form a single, *combined, units permutation*. So the randomized layout can be obtained by forming a table with (i) Rows and Columns numbers in the permuted order for the combined, units permutation and (ii) Varieties in systematic order for an LSqD, say A–E in the first row and then these letters just cycled one position to the right from one row to the next. Re-ordering the rows of this table so that Rows and Columns are in standard order produces the randomized design in Supplement Figure 1.

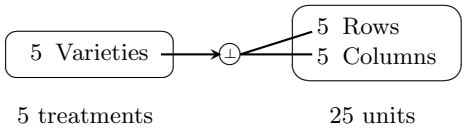

FIGURE 2 Factor-allocation diagram showing the treatments allocation to units for a Latin square design: the arrow indicates that the allocation is randomized; the '⊥' at the end of the arrow indicates that an orthogonal design is used; the two lines from the '⊥' indicate that the Treatments are allocated to the combinations of Rows and Columns using the design.

The factor-allocation diagram. Figure 1 suggests that the allocations are exhibited in a *factor-allocation diagram*. Figure 2 does this for the LSqD. It consists of two *panels*, each of which contains the factors that index the, possibly conceptual, objects that label the panels. In this paper, we always label the objects as treatments and units; lower case labels are used for objects to distinguish them from factors whose names use an initial capital letter. There is one *allocation*: the allocation of treatments to units using randomization, as indicated by the solid arrow. The factors in different panels differ in their roles in the allocations: (i) the factor in the treatments panel is allocated; and (ii) the factors in the units panel are recipient factors. In this case, the permutations used in randomizing the design determine that Rows and Columns are crossed, and so there is no nesting of factors in the units panel.

The initial allocation model. In keeping with Figure 1, the *initial allocation model* is now derived from the factor-allocation diagram. Terms are formed by taking all combinations of the factors within a panel, including individual factors and subject to the restriction that a factor that is nested cannot occur without its nesting factors. Each *term* is a combination of factors, whose levels are the observed combinations of the levels of the constituent factors. In the initial allocation model, terms with factors that were only ever allocated are designated as fixed. The other terms are designated as random and each has a variance component associated with it. The initial allocation model for this example is

the same as in (1). Thus, the crossing of Rows and Columns inherent to the anticipated model has been reflected in the combined, units permutation employed in allocating treatments to units for the chosen design. This permutation has determined the factor-allocation diagram, which was then used to produce the initial allocation model. The desired outcome of the crossing and nesting encapsulated in the initial allocation model matching that inherent to the anticipated model has been achieved.

The homogeneous allocation model. The *homogeneous allocation model* is obtained from the initial allocation model, as portrayed in Figure 1, by transferring terms from the fixed part to the random part of the model, or vice-versa, and by adding interactions between terms from different panels, or block-treatment interactions. Our view on the controversial topic of deciding whether a term is fixed or random is that, in essence, it is a decision about the appropriate model for the term. To be *random*, the effects for a term need to be behave in a manner that it is appropriate to characterize them by a probability distribution, by default a normal distribution. Otherwise, they should be designated as *fixed*, which allows for arbitrary values of the effects and so involves less restrictive assumptions. In practice, a term is likely to be declared (i) random if its effects exhibit random behaviour and there is, at least conceptually, a large number of population levels for its factors, and (ii) fixed if there is either a small or large number of population levels and its effects exhibit systematic or other non-random behaviour. A further consideration for declaring a term to be random is whether it is desired to draw conclusions using it beyond the observed levels of its factors, such as for a units term that has treatment information confounded with it and conclusions about the treatments beyond the observed levels of the units terms are wanted, or conclusions about the future performance of a set of its levels are required, such as when a subset of plant or animal genotypes is to be selected from a candidate set of observed genotypes. Notwithstanding this, the behaviour of the effects for such a term need to conform to the behaviour required for a random term and it is incumbent on the designer to ensure that the number and levels of the factors involved ensures this. A commonly advanced criterion is that for declaring a term to be random is that it includes at least one factor whose levels are randomly sampled, or could be so regarded. While this may be a desirable property, it appears not to be a sufficient criterion. For example, it would not be valid to treat as random, a term for Blocks that has been the randomly sampled along a field fertility or laboratory temperature trend. The crucial point is that random sampling of the levels does not guarantee an appropriate distribution of the effects.

If it were decided that it is appropriate to assume that Rows and Columns are fixed in this example and that there is block-treatment additivity, then the homogeneous allocation model is:

$$\text{Mean} + \text{V} + \text{R} + \text{C} \quad | \quad \text{Mean} + \underline{\text{R:C}}. \tag{3}$$

It is a *fixed model*, having only the Mean and an identity term being random. Suppose that ultimately it was decided that this change is inappropriate and the homogeneous allocation model is given by (1), which is equivalent to a randomization model.

Anatomy table. An important aspect of the approach outlined in Figure 1 is the production of an anatomy table for the homogeneous allocation model as a means of checking and understanding the design by studying the confounding integral to it. Brien (2017) describes the *anatomy* as the structure built on the allocations in the design and is expressed in the confounding relationships between the sources derived from the terms in a model for the design. A *source* for a term is the subspace of the units space for the term made orthogonal to all column spaces of the indicator matrices for marginal terms from the panel. The *units space* is the column space for an identity term. A term is *marginal* to another term from the same panel if the column space of its indicator matrix is a subspace of the column space of the other term's indicator matrix. For the units panel, the Mean term is marginal to all terms from that panel and the Rows and Columns terms are marginal to the term R:C. The degrees of freedom (DF) for a source is the dimension of its subspace. Sources specify contrasts on the units, those involving single factors being referred to as *main effects* and those involving more than one factor specify *interactions* and/or *nested effects*, depending on the crossing and nesting relations between the factors intrinsic to the factor allocation for the design. To show the nature of a source, a '#' separates any factors that interact, these being the factors that do not nest any factor; all the factors that nest another factor, if any, are enclosed within square braces ('[...]'), being separated by ':' to indicate a combined factor whose levels are the observed combinations of the nesting factors.

The *anatomy table* is like a skeleton ANOVA table, being similar in form to the multistrata skeleton ANOVA tables produced by Genstat (VSN International, 2022b). However, anatomy and ANOVA tables differ in their function. The sole purpose of anatomy tables is to reveal the confounding and

aliasing relationships between the terms associated with a design, thus facilitating a study of the design's properties. These relationships are ascertained from an analysis of the projection matrices derived from the terms. The role of skeleton ANOVA tables has generally been to establish the form of the analysis of the data from a designed experiment prior to performing an ANOVA on the data. This role is no longer required when linear mixed model fitting is employed, whereas anatomy tables remain relevant because their focus is on the design.

The confounding in a design is summarized in an anatomy table that is divided into three major columns, one for the units sources and DF, a second for the treatments sources and DF and, if possible, a third for the *expected mean square* (EMS) for each of its rows (e.g. Table 1a). The sources in each Source column correspond to the terms in the homogeneous allocation model derived from one of the panels in the factor-allocation diagram, although the treatments panel may at times involve both treatments and units factors. To manually construct the anatomy table, first list the units sources and their DF, thereby creating the first major column. Then create the second major column by adding each treatment source and its DF, aligning it with the units source(s) with which the treatments source is confounded, putting multiple treatment sources confounded with the same units source one under the other. For a main-effects treatments source, the source with which it is confounded is usually the units source to which it was allocated. The confounding of interaction treatments sources may depend on the design, but often an interaction source is confounded with the units source that occurs lowest in the table among the units sources with which main-effects sources for the factors in the interaction term are confounded. Once all treatments sources have been placed in the table, a Residual line is added to any units source whose DF in the first major column exceeds the sum of the treatments sources confounded with it; the difference in these two DFs is the DF for the Residual source. Thus the alignment of the treatment sources with units source displays the confounding relationships, the units sources being partitioned by the treatment sources. The two sets of sources, with their DF, specify the joint decomposition of the units space. Clearly, the division of the factors into two panels, as illustrated in the factor-allocation diagram, is essential to representing the confounding.

TABLE 1 The anatomy and traditional, skeleton ANOVA tables under the homogeneous allocation model for an experiment on a 5×5 grid of plots where both row and column effects are anticipated: R = Rows; C = Columns; V = Varieties; EMS = Expected Mean Square; DF = Degrees of Freedom.

(a) Anatomy table.

units		**treatments**		**EMS**†				
Source	**DF**	**Source**	**DF**	σ^2_{RC}	σ^2_{C}	σ^2_{R}	σ^2_{μ}	$\theta_{\cdot}$
Mean	1	Mean	1	1	5	5	25	θ_{μ}
Rows	4			1		5		
Columns	4			1	5			
R # C	16	Varieties	4	1				θ_{V}
		Residual	12	1				
Total	25							

(b) Traditional, skeleton ANOVA table.

		EMS†				
Source	**DF**	σ^2	σ^2_{C}	σ^2_{R}	σ^2_{μ}	$\theta_{\cdot}$
Mean	1	1	5	5	25	θ_{μ}
Rows	4	1		5		
Columns	4	1	5			
Varieties	4	1				θ_{V}
Residual	12	1				
Total	25					

†Each $\sigma^2_{\cdot}$ is a variance component whose subscripts are comprised of the first letter of each factor in the corresponding term and the numbers in the table are the coefficients of the variance components in the EMSs. Each $\theta_{\cdot}$ is a quadratic form in the expectation of a response variable, with the subscript indicating the source whose scaled projection matrix is the matrix of the quadratic form. Here, θ_{V} simplifies to the sum of the squared population Variety effects, multiplied by five, the number of Varieties replicates, and divided by four, the Varieties DF

The anatomy table, without EMS, can be obtained using the function `designAnatomy` from the `R` package `dae` (see the `R` script in Supplement Section C.1). The function uses a formula for the terms in each panel and a design or layout in the form of an allocation of the treatments factors to the unit factors. To determine the confounding, it forms projection matrices for all sources, and then uses them to check, for each treatments source, with which of the units sources it is confounded. To check if a treatments source is confounded with a units source, the harmonic mean of the canonical efficiency

factors, the *A-efficiency criterion*, for these two sources is obtained (John and Williams, 1995); the A-efficiency criterion can be interpreted as giving the proportion of the information for the treatment source associated with the units source. The canonical efficiency factors are the nonzero eigenvalues of a certain product of the projection matrices for the two sources in each pair. The number of canonical efficiency factors is equal to the DF of the confounded source (for details see Supplement Section D). For each treatments source in an orthogonal design, all canonical efficiency factors are one for just one units sources. Consequently, all A-efficiency criteria are one, the maximum value for an A-efficiency criterion. A design is said to be *A-optimal* for a fixed model if its A-efficiency criterion is maximized for each treatments source for the smallest units entity with which the source is confounded. It may also be A-optimal for a mixed model. All of the designs presented in this paper are A-optimal for a mixed model and are universally optimal for a fixed model i.e. they are A-, D- and E-optimal (Shah and Sinha, 1989). The EMSs in the anatomy table are obtained using the rules in the Appendix.

Table 1a shows the anatomy table for the LSqD; Examples 1 and 2 in the Appendix describe the use of the rules to obtain the EMS for the Rows and R # C lines in this table. Here, the Varieties source was checked for confounding with the sources Mean, Rows, Columns and R # C. It was found that, only for Varieties and R # C, were there canonical efficiency factors, there being four of them all of which were equal to one. Thus, the A-efficiency for Varieties is one and all Variety information is confounded with R # C. The LSqD design is orthogonal and so is A-optimal under the anticipated model in (1). It is also seen that, of the total 16 DF for R # C, 12 are left over for estimating σ^2_{RC}.

All designs involve confounding, but this is not widely recognized. This is to be expected as traditional ANOVA tables have only a single sequence of sources listed in the order in which they are to be fitted and so there is no opportunity to portray the confounding (cf. the anatomy table with two sets of sources). Also contributing to this neglect of confounding is the general practice of not accommodating specific sources for all recipient terms in the table; in this case, R # C is habitually omitted. To demonstrate how the tables differ, the traditional, skeleton ANOVA table is presented in Table 1b. It is based on the model (see for example Littell et al., 2006):

$$\mathrm{Mean} + \mathrm{V} \quad | \quad \mathrm{Mean} + \mathrm{R} + \mathrm{C} + \underline{\mathrm{units}}, \tag{4}$$

where units indicates a term for each observation whose effects are assumed to be independent and identically distributed with common variance σ^2.

The decompositions of the units space portrayed in Tables 1a and 1b are equivalent. The crucial difference is that, while one can see that the Varieties source is confounded with the interaction R # C in Table 1a, this confounding is not apparent in Table 1b. Certainly, the EMSs in Table 1b show that the Varieties source is only affected by Residual variation, the overall differences between Rows and Columns having been eliminated. However, the origins of this variation are not obvious. On the other hand, Table 1a suggests that the Varieties source is confounded with unit variation, rather than block-treatment interaction. We will return to this point in Section 3.

Table 1a includes the source Mean for the overall mean, μ. It is often omitted from ANOVA tables, but it is retained here for completeness and to explore the ramifications of the two Mean sources in the models. However, because both a fixed and a random effect are included for μ in the model for the experiment, the variance component σ^2_μ and the fixed contribution θ_μ occur in the EMS for the source. The EMS for the Mean row shows that the two parameters are confounded and the value of μ, incorporated in θ_μ, is inseparable from the variability in μ, σ^2_μ. This is the case for all comparative experiments, so called, because only treatments contrasts are estimable.

2.2 No Columns effects expected

In this section, an examination is made of the effects on the design and analysis of experiments on a grid when no Columns effects are expected.

The anticipated model. Speculating that there will only be Rows effects and that their effects are likely to follow a systematic trend, it is concluded that the anticipated model should include a fixed Rows term. Thus the symbolic anticipated model is:

$$\mathrm{Mean} + \mathrm{R} + \mathrm{V} \quad | \quad \mathrm{Mean} + \underline{\mathrm{R{:}C}}. \tag{5}$$

That there is a Rows main effect in (5), but not a Columns main effect, means that it is inherent to this model that Columns is *nested* within Rows. Here, as opposed to the example in Section 2.1, the crossing

and nesting inherent in the anticipated model for the current example does not match the crossing and nesting innate to the physical site.

Design selection & construction. A design for which Columns are nested within Rows is required. An RCBD is the A-optimal design under the anticipated model; an LSqD may be less efficient. A strategy for avoiding an efficiency loss might be to always employ an LSqD and to omit from the analysis any block terms deemed to be unimportant. However, this is a sometimes-pool strategy (Janky, 2000), a practice that Janky argues is not applicable if a randomization argument is being used to justify the assumption of random blocking effects; even if this argument is not being relied upon, Janky concludes the strategy should not be used routinely because (i) it compromises the size of the Type I error rate, and (ii) there can be power losses, and power gains that, when they occur, are generally modest.

The appropriate permutations for an RCBD are to permute (i) the Rows and (ii) the Columns within Rows. An `R` script to produce and plot a randomized layout is in Supplement Section C.2 and the plotted layout is in Supplement Figure 2. The permutation chosen for Rows was (4,2,1,3,5) and for the Columns of each Row was: (2,4,5,3,1), (2,5,1,4,3), (3,4,1,5,2), (4,3,2,5,1), (4,3,2,1,5). These are combined to form a single, units permutation.

The factor-allocation diagram. Figure 3 exhibits the factor allocation for the RCBD. The nesting in this diagram reflects the permutations used in randomizing the design.

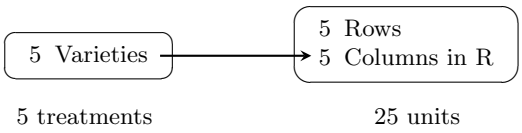

FIGURE 3 Factor-allocation diagram showing the treatments allocation to units for a randomized complete-block design: the arrow indicates that Varieties are allocated to Columns within R using randomization; R = Rows.

The initial allocation model. The initial allocation model for this example is:

$$\text{Mean} + \text{V} \quad | \quad \text{Mean} + \text{R} + \underline{\text{R:C}}. \tag{6}$$

Again, our design selection has been successful because the crossing and nesting encapsulated in the initial allocation model, that reflects the permutations for the chosen design, matches the crossing and nesting inherent in the anticipated model.

The homogeneous allocation model. The initial allocation and anticipated models differ in whether Rows are random or fixed. Also, suppose that it is appropriate to assume that there is block-treatment additivity, as was done for the anticipated model. Thus, the initial allocation model needs to be changed to produce a homogeneous allocation model that is the same as the anticipated model in (5).

Anatomy table. Table 2 has the anatomy table (an `R` script that produces the table, without EMSs, is in Supplement Section C.2). It shows that (i) Varieties is confounded with differences between Columns within Rows, (ii) the design is orthogonal and, as the anticipated model is a fixed model, it is A-optimal, and (iii) there are 16 Residual DF for estimating σ^2_{RC}.

TABLE 2 Anatomy table under the homogeneous allocation model for an experiment on a 5×5 grid of plots where no column effects are anticipated: R = Rows; V = Varieties; EMS = Expected Mean Square; DF = Degrees of Freedom.

units		**treatments**		**EMS**		
Source	**DF**	**Source**	**DF**	σ^2_{RC}	σ^2_{μ}	$\theta.$
Mean	1	Mean	1	1	25	θ_{μ}
Rows	4			1		θ_{R}
Columns [R]	20	Varieties	4	1		θ_{V}
		Residual	16	1		
Total	25					

3 Block-treatment and unit-treatment interactions

In Section 2, the anticipated models assumed both no block-treatment and no unit-treatment interactions, where *block-treatment interactions* result from the differences between treatments not being the same in different blocks and *unit-treatment interactions* arise when differences between treatments, when allocated to the same unit if that was possible, are not the same for different units (Hinkelmann and Kempthorne, 2008, Section 9.6). In general, it is impossible to detect unit-treatment interactions, but there exist circumstances in which block-treatment interactions need to be included in the model, as is envisaged in the modification of the initial allocation model to form the homogeneous allocation model. Indeed some commonly used approaches to model building for comparative experiments essentially incorporate all possible block-treatment interactions (Littell et al., 2006; Loughin, 2005; Bate and Chatfield, 2016a,b). In this section, situations that include block-treatment interactions are investigated.

3.1 The randomized complete-block design

Suppose that, during the discussion with the researcher about the design of the experiment on a 5×5 grid of plots, it had been suggested that, not only were no overall column differences expected, but that Varieties may interact with Rows. How does the factor-allocation approach deal with this situation?

The anticipated model. Presuming that the block-treatment interactions can be assumed to be random, the symbolic anticipated model is now:

$$\text{Mean} + \text{R} + \text{V} \quad | \quad \text{Mean} + \underline{\text{R:V}} + \underline{\text{R:C}}. \tag{7}$$

This model involves a random two-factor term, R:V, whose main-effect terms are fixed. As Harville (1991) suggests, this would be inappropriate if the criterion for declaring a term to be random is that the levels of one or more of its factors must be randomly sampled. If Rows and Varieties are not randomly sampled then this criterion mandates that R and V, and hence the R:V interaction, are fixed. If Rows can be regarded as being a random sample, then this criterion would result in both Rows and the R:V interaction being random. However, the criterion that we have espoused in Section 2.1 is that a term will be random only if its effects can legitimately be viewed as conforming to a probability distribution. For this criterion, each term's effects are considered separately. Thus, taking the Rows main effects, is it reasonable to assume that a probability distribution is an appropriate model for them? Suppose there is a systematic trend across the Rows, in which case Rows must be fixed, regardless of whether they are randomly sampled. Turning to the R:V interaction effects, are these random across the Rows-Columns combinations? If yes, a probability distribution is a feasible model and so the interaction can be regarded as a random term, irrespective of whether the Rows and Varieties effects are fixed or random. If not, R:V should be designated as fixed.

Design selection & construction. An RCBD remains the proposed design.

The factor-allocation diagram. Figure 3 shows the factor allocation for the RCBD.

The initial allocation model. Again, the initial allocation model is (6).

The homogeneous allocation model. To produce the homogeneous allocation model from the initial allocation model, Rows needs to be designated as a fixed term and the random term R:V needs to be added. The resulting homogeneous allocation model is then the same as the anticipated model in (7).

Anatomy table. The anatomy table is in Table 3a (an `R` script that produces the table, without EMSs, is in Supplement Section C.2). It can be obtained from Table 2 by (i) labelling the second major column as treatments-blocks, (ii) adding the block-treatment interaction source R # V to the treatments sources and adding a variance component for it (σ^2_{RV}), and (iii) replacing the Rows variance component with a fixed contribution θ_{R} to reflect that Rows are now regarded as fixed. As in Table 2, Varieties is confounded with differences between Columns within Rows in Table 3a. In addition, R # V is also confounded with Columns [R]. Indeed, Varieties and R # V together *exhaustively confound* the units source Columns [R], as evidenced by the lack of a Residual for Columns [R]. The EMSs show that variability between Columns within Rows (σ^2_{RC}) cannot be separated from the interaction variance of Rows and Varieties (σ^2_{RV}), because all EMSs include both components. Nonetheless, Varieties can be tested using the R # V mean square.

The effect of the exhaustive confounding of Columns [R] is that the variance matrix under the proposed homogeneous allocation model is singular. Thus, the homogeneous allocation model cannot be fitted with mixed model fitting software. An equivalent nonsingular model that can be fitted must

TABLE 3 Anatomy table for the homogeneous allocation model for an RCBD with block-treatment interactions: R = Rows; V = Varieties; EMS = Expected Mean Square; DF = Degrees of Freedom.

(a) Anatomy table.

units		treatments-blocks		EMS			
Source	**DF**	**Source**	**DF**	σ^2_{RC}	σ^2_{μ}	σ^2_{RV}	$\theta.$
Mean	1	Mean	1	1	25	1	θ_{μ}
Rows	4			1		1	θ_{R}
Columns [R]	20	Varieties	4	1		1	θ_{V}
		R # V	16	1		1	
Total	25						

(b) Traditional, skeleton ANOVA table.

		EMS†		
Source	**DF**	σ^2_{RV}	σ^2_{μ}	$\theta.$
Mean	1	1	25	θ_{μ}
Rows	4	1		θ_{R}
Varieties	4	1		θ_{V}
R # V	16	1		
Total	25			

be identified. This can be achieved by removing either of the terms R:C or R:V, resulting in what Brien and Demétrio (2009) term a *model of convenience*: a model that does not include all the possible terms at play in the experiment, but which is nonsingular and can be fitted. We would remove the block-treatment interaction term R:V, so that the term germane to the allocations used in the design is retained. On the other hand, Piepho et al. (2003), who give a method for formulating linear models related to ours, appear to favour removing R:C. However, which is removed makes no difference to the fit of the model. The important point is that, whichever term is retained, the estimated value of its variance component is the sum of the confounded components. Thus, an estimate of σ^2_{RC} is in reality an estimate of $\sigma^2_{\mathrm{RC}} + \sigma^2_{\mathrm{RV}}$. The *nonsingular homogeneous allocation model*, derived from the homogeneous allocation model in (7) by removing the term R:V, is the same as the anticipated and initial allocation models in (5), for which no block-treatment interaction was assumed.

The traditional approach to formulating an ANOVA for an RCBD, as expounded by Littell et al. (2006) and Loughin (2005), does separate the blocking and treatment aspects of a design. Crucially, it does this by identifying factors that are related to blocks in the design, rather than the full set of units factors. For the example, it would only nominate Rows as a blocking factor and Varieties as a treatment factor, omitting Columns. Based on this, it seeks to identify *experimental units* as potential sources of variation in the experiment, where an 'experimental unit is defined as the smallest entity to which a treatment is independently applied' (Littell et al., 2006, Section 4.2). In practice, their experimental unit is the combination of the block and treatment factors to which the levels of the treatment factor are assigned. For the example, their experimental unit is said to be a Rows-Varieties combination. This appears to be a contradiction in terms in that 'each treatment must be assigned to an experimental unit' (Littell et al., 2006, Section 4.2) and one cannot allocate Varieties to Rows-Varieties combinations because the Variety associated with one of these combinations cannot be changed. Rather, Varieties are randomly allocated to Rows-Columns combinations and this determines the Rows-Varieties combination assigned to each Rows-Columns combination. Thus, it is in fact the Rows-Columns combinations that are the experimental units, as depicted in Figure 3.

The traditional, skeleton ANOVA table, which omits the source Columns [R], is given in Table 3b. Thus, the inability of a traditional table to portray confounding gives the impression here that Varieties is tested against just the block-treatment interaction, obscuring the contribution of Columns [R] to the test for Varieties. On the other hand, Table 3a makes explicit that both the *inherent random variability* between Columns (units) within Rows (blocks) and the interaction between Rows and Varieties (a block-treatment interaction) contribute to the test for Varieties. The proposed paradigm addresses the distinction that Samuels et al. (1991) made between inherent random variability and block-treatment interaction by including separate terms for each, as Harville (1991) suggests.

3.2 The generalized randomized-block design

Addelman (1969) recommended the increased use of the generalized randomized-block design (GRBD), an important reason being that it allows testing for block-treatment interactions.

A glasshouse experiment is being planned to compare the effects of four Zinc levels on plants of a medic species. It is to involve 12 replicates and the experimental area can accommodate 48 pots in grid of four lanes by 12 positions, each pot having a single plant.

The anticipated model. Previous experience is that the differences between pots in the same lane separated by more than two pots are likely to be larger than between those separated by no more than two pots. Also, pots in the front pair of lanes are likely to differ from pots in the back pair of lanes. That is, Blocks consisting of eight pots arranged in two lanes by four positions are likely to be relatively homogeneous. Further, suppose that it is thought that the response to Zinc may differ between the Blocks. Thus, the factors for the experiment are Blocks, Pots and Zinc and the anticipated model is:

$$\text{Mean} + \text{Z} \quad | \quad \text{Mean} + \text{B} + \text{B:Z} + \underline{\text{B:P}}. \tag{8}$$

Design selection & construction. A GRBD is the A-optimal design. The appropriate permutations for a GRBD are to permute (i) the Blocks and (ii) the Pots within Blocks. An `R` script to produce and plot a randomized layout is in Supplement Section C.3 and the plotted layout is in Supplement Figure 3.

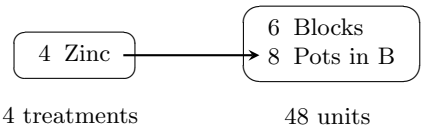

FIGURE 4 Factor-allocation diagram showing the treatments allocation to units for a generalized randomized block design: the arrow indicates that Zinc is allocated to Pots within B using randomization; B = Blocks.

The factor-allocation diagram. Figure 4 exhibits the factor allocation for the GRBD. This example raises a problem for specifying the experimental unit as the interaction of block and treatment factors, as in Littell et al. (2006). Here the experimental units are not the combinations of Blocks and Zinc; they are the combinations of Pots and Blocks, but Pots is traditionally omitted because, while it is a units factor, it is not a block factor.

The initial allocation model. The initial allocation model for this example is:

$$\text{Mean} + \text{Z} \quad | \quad \text{Mean} + \text{B} + \underline{\text{B:P}}, \tag{9}$$

The homogeneous allocation model. The initial allocation model does not contain the term for the block-treatment interaction that is in the anticipated model. Thus, to form the homogeneous allocation model, this term would be added to the initial allocation model so that the resulting model is the same as the anticipated model in (8).

Anatomy table. Table 4 has the anatomy table (an `R` script that produces the table, without EMSs, is in Supplement Section C.3). Using the rules to add the EMSs to the table is illustrated in Supplement Section E.3. The table shows (i) the design is orthogonal and so is A-optimal and (ii) the sources Zinc and B # Z are confounded with Pots differences within Blocks. It also reveals that the denominator in an F-test for Zinc is the B # Z mean square, with 15 DF, rather than the Residual mean square with 24 DF. The block-treatment interaction B # Z is tested using the Residual mean square.

4 Further insights provided by considering the factor allocation

Three examples in which time is a factor are used to illustrate how similar factors can be either allocated or recipient factors, that recipient factors can be of interest to the researcher, that systematic allocation can be accommodated, and that pseudoreplication causes unwelcome confounding: (i) an experiment in which measurements at different times are made on different units; (ii) a longitudinal experiment in which repeated measurements are made on the same entity, and (iii) a repeated-measurements experiment in which the different treatments are applied to the same entity at different times.

TABLE 4 Anatomy table for the homogeneous allocation model for a GRBD with block-treatment interactions: B = Blocks; Z = Zinc; EMS = Expected Mean Square; DF = Degrees of Freedom.

units		treatments-blocks		EMS				
Source	DF	Source	DF	σ^2_{BP}	σ^2_{B}	σ^2_{μ}	σ^2_{BZ}	$\theta.$
Mean	1	Mean	1	1	8	48	2	θ_{μ}
Blocks	5			1	8		2	
Pots [B]	42	Zinc	3	1			2	θ_{Z}
		B # Z	15	1			2	
		Residual	24	1				
Total	48							

4.1 An experiment with a randomized time factor

A design is required for a conventional greenhouse experiment to investigate Zinc effects on plants of a medic species. A response over five weeks is to be measured, but the measurement of the response requires destructive harvesting of the plants. It is to involve four levels of Zinc and there are to be eight replicates of the Zinc-Weeks combinations.

The anticipated model. The total number of units is 160 and pots must be arranged four deep on a long bench so that the experimental area is four Lanes by 40 Positions. In discussions with the researcher, it is decided to maximize the precision of response differences between the Zinc levels, the response differences between weeks being expected to be large. Because it is known that there are differences between the front and back pairs of Lanes, it is best if a replicate of the Zinc treatments is kept within a pair of Lanes. This implies a complete block that is two Lanes by 10 Positions. It is felt that differences between pots at opposite ends of such a block are likely to be larger than between pots at the same end and so subdividing a block to form main units is seen as advantageous. Thus, the factors for the experiment are Blocks, MainUnits, Pots, Zinc and Weeks. The anticipated model is:

$$\text{Mean} + \text{Z} + \text{W} + \text{Z:W} \quad | \quad \text{Mean} + \text{B} + \text{B:M} + \underline{\text{B:M:P}}. \tag{10}$$

Design selection & construction. An orthogonal split-unit design is A-optimal under the anticipated model for Zinc and for Weeks, given that Weeks is allocated to MainUnits within Blocks. To achieve the desired precision for the comparison of Zinc levels, a main unit could be formed from four pots in a 2×2 grid and the Zinc levels randomized to the pots. Five main units are needed in each of eight blocks to accommodate the five weeks of destructive observations, with four blocks in each pair of Lanes. The weeks are randomized to the main units. To randomize this design, permute (i) Blocks, (ii) MainUnits within Blocks, and (iii) Pots within each combination of Blocks and MainUnits. An `R` script to produce and plot a randomized layout is in Supplement Section C.4 and the plotted layout is in Supplement Figure 4. A combined, units permutation, that combines the three permutations for this design, is applied as a single randomization by `designRandomize` from `dae`.

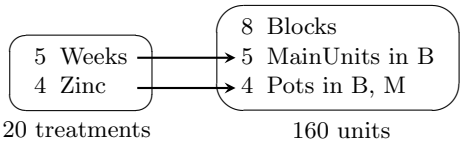

FIGURE 5 Factor-allocation diagram showing the treatments allocation to units for the split-unit experiment that randomizes the time factor Weeks; the arrows indicate that randomization is used to allocate Weeks to MainUnits within B and Zinc to Pots within each combination of B and M; B = Blocks, M = MainUnits.

The factor-allocation diagram. Figure 5 exhibits the factor allocation.

The initial allocation model. The initial allocation model for this example is the same as the anticipated model in (10).

The homogeneous allocation model. Here a policy of including all possible block-treatment interactions is followed and these need to be added to the initial allocation model to yield the following homogeneous allocation model:

$$\text{Mean} + \text{Z} + \text{W} + \text{Z:W} \quad | \quad \text{Mean} + \text{B} + \text{B:Z} + \text{B:M} + \text{B:W} + \underline{\text{B:Z:W}} + \underline{\text{B:M:P}}. \tag{11}$$

TABLE 5 Anatomy table for the homogeneous allocation model for a longitudinal experiment in which Weeks is randomized: B = Blocks; M = MainUnits; W = Weeks; Z = Zinc; W = Weeks; EMS = Expected Mean Square; DF = Degrees of Freedom.

units		treatments-blocks		EMS							
Source	DF	Source	DF	σ^2_{BMP}	σ^2_{BM}	σ^2_{B}	σ^2_{μ}	σ^2_{BZW}	σ^2_{BZ}	σ^2_{BW}	$\theta_{\cdot}$
Mean	1	Mean	1	1	4	20	160	1	5	4	θ_{μ}
Blocks	7			1	4	20		1	5	4	
MainUnits [B]	32	Weeks	4	1	4			1		4	θ_{W}
		B # W	28	1	4			1		4	
Pots [B:M]	120	Zinc	3	1				1	5		θ_{Z}
		Z # W	12	1				1			θ_{ZW}
		B # Z	21	1				1	5		
		B # Z # W	84	1				1			
Total	160										

Anatomy table. The anatomy table, without EMSs, was produced using the `R` script in Supplement Section C.4 and the EMSs derived using the rules in the Appendix, resulting in Table 5. It shows that the sources Weeks and B # W are confounded with MainUnits [B]. The remaining treatments sources are confounded with Pots [B:M]. It also shows that the only variance components that are separately estimable are those for Blocks and B:Z. To fit the homogeneous allocation model, terms would need to be removed, for example, B # W and B # Z # W. While hypothesis tests for all three fixed effects are possible, the F-test for the Zinc main effect would use the B # Z mean square, which has 21 DF, as opposed to the B # Z # W mean square with 84 DF. If block-treatment additivity is assumed then the tests for Zinc and Z # W would be against a Residual mean square with 105 DF. The model under this assumption, the initial allocation model in (10), is equivalent to a randomization model and a randomization analysis would be available.

4.2 A factor of interest that is not allocated: a longitudinal experiment

It often happens that the factors of interest are the treatment factors that have been allocated in the design for an experiment. However, it is common for factors intrinsic to the units, and so (potential) recipient factors, to be of interest also. Some examples are Times of observation in longitudinal studies, the Sex of animals in animal studies and Tasters in sensory evaluation experiments. These factors block the units, and interactions between them and treatment factors are block-treatment interactions that are often of interest to researchers.

Consider an experiment in a glasshouse that has equipment to automatically image plants daily, the images being processed to produce a measure related to plant biomass. Suppose that the experiment to investigate the effects of four levels of Zinc on medic plants, discussed in Section 3.2, is to be run in this glasshouse and that the plants are to be imaged over 14 Days.

The anticipated model. The factors for this experiment are Blocks, Pots, Zinc and Days. The anticipated model for this experiment extends the anticipated model for the GRBD in Section 3.2 by adding terms for the longitudinal factor Days and all interactions of the GRBD terms with Days. Here, Days is acting as a blocking factor by grouping the units according to when they were produced. Experience with such experiments suggests that smooth trends over the Days and random deviations from these trends can be expected for some terms involving Days; it is also thought possible that there will be

correlations between Days and that Days variances will be heterogeneous. The anticipated model is:

$$\text{Mean} + \text{Z} + \texttt{td}(\text{D}) + \text{Z}{:}\texttt{td}(\text{D}) \quad | \quad \begin{array}{l}\text{Mean} + \text{B} + \texttt{dev}(\text{D}) + \text{Z}{:}\texttt{dev}(\text{D}) + \text{B}{:}\texttt{td}(\text{D}) + \text{B}{:}\texttt{dev}(\text{D}) + \\ \text{B}{:}\text{Z} + \text{B}{:}\text{P} + \text{B}{:}\text{Z}{:}\text{D} + \underline{\text{B}{:}\text{P}{:}\texttt{sch}(\text{D})},\end{array} \tag{12}$$

where `td(D)` indicates a smooth trend over Days, `dev(D)` indicates random deviations from a smooth trend over Days and `sch`(D) allows for serial correlation (sc) between Days and for heterogeneous (h) Days variances. The form of the smooth trend and the correlation between Days are left unspecified.

Design selection & construction. It was established, in Section 3.2, that a GRBD is the A-optimal design for allocating Zinc to Pots. An `R` script to produce and plot a randomized layout is in Supplement Section C.5 and the plotted layout is in Supplement Figure 5.

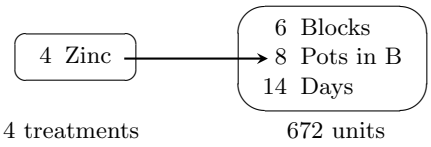

FIGURE 6 Factor-allocation diagram showing the treatments allocation to units for the longitudinal experiment that uses a generalized randomized block design: the arrow indicates that Zinc is allocated to Pots within B using randomization; B = Blocks.

The factor-allocation diagram. Figure 6 exhibits the factor allocation for the longitudinal experiment. Of note is that Days is a recipient factor, not being allocated; it forms 'blocks' of units and is intrinsic to the units, the units being pots-days combinations. While these experiments have been referred to as split-units-in-time (e.g. Steel and Torrie, 1960), they are not true split-unit designs because Days is not allocated, being a recipient factor, and Days are not nested within Pots. This is in marked contrast to the experiment in Section 4.1 where the time factor Weeks is randomized, and so an allocated factor, and the subunits factor Pots is nested within MainUnits. However, like Week,s and in spite of being a units factor, Days is of interest because growth trends for the different Zinc levels are to be compared. In conclusion, because of the very different allocations for this type of experiment compared to the split-unit design, we see split-unit-in-time as a misnomer that serves to obscure several of its features, which will be elaborated upon below.

The initial allocation model. The initial allocation model for this example is:

$$\text{Mean} + \text{Z} \quad | \quad \text{Mean} + \text{B} + \text{B}{:}\text{P} + \text{D} + \text{B}{:}\text{D} + \underline{\text{B}{:}\text{P}{:}\text{D}}, \tag{13}$$

This model differs from the initial allocation model for a split-unit design given in (10) in that it includes terms that have Days combined with other units factors, but not Days combined with the treatments factor Zinc. In particular, the term B:D is automatically included because of the relationships between the units factors; it is not always included in the model for the split-unit-in-time, as it should be (Piepho et al., 2004). The interaction terms Z:D and B:Z:D are not included here because they are block-treatment interactions in that they involve both treatments and units factors.

The homogeneous allocation model. The homogeneous allocation model, obtained from the initial allocation in (13) by moving Days to the expectation model and adding the terms for the block-treatment interactions that are in the anticipated model, is

$$\text{Mean} + \text{Z} + \text{D} + \text{Z}{:}\text{D} \quad | \quad \text{Mean} + \text{B} + \text{B}{:}\text{Z} + \text{B}{:}\text{P} + \text{B}{:}\text{D} + \text{B}{:}\text{Z}{:}\text{D} + \underline{\text{B}{:}\text{P}{:}\text{D}}. \tag{14}$$

That is, the only difference between a term in this model and the term with the same factors in the anticipated model is in their parameterizations. Restricting the homogeneous model to the assumption of uniform covariance for all of its terms retains the possibility of forming an anatomy table incorporating sources for all of its terms. Of course, the anticipated model would provide a good initial model for the analysis of observed responses.

Anatomy table. The `R` script in Supplement Section C.5 was used to produce the anatomy table, without EMSs, and the EMSs derived using the rules in the Appendix, yielding Table 6. As in Table 4, it shows that the Zinc and B # Z are confounded with Pots [B]. It also shows that Z # D and B # Z # D are confounded with D # P [B] and that Zinc and Z # D can be tested against the Residual for D # P [B]. What this table does show is that, like the anatomy of a split-unit design with block-treatment additivity,

TABLE 6 Anatomy table for the homogeneous allocation model for a longitudinal experiment that uses a GRBD: B = Blocks; P = Pots; D = Days; Z = Zinc; D = Days; EMS = Expected Mean Square; DF = Degrees of Freedom.

units		treatments-blocks		EMS							
Source	DF	Source	DF	σ^2_{BPD}	σ^2_{BD}	σ^2_{BP}	σ^2_{B}	σ^2_{μ}	σ^2_{BZD}	σ^2_{BZ}	$\theta_{\cdot}$
Mean	1	Mean	1	1	8	14	112	672	2	28	θ_{μ}
Blocks	5			1	8	14	112		2	28	
Pots [B]	42	Zinc	3	1		14			2	28	θ_{Z}
		B # Z	15	1		14			2	28	
		Residual	24	1		14					
Days	13			1	8				2		θ_{D}
B # D	65			1	8				2		
D # P [B]	546	Z # D	39	1					2		θ_{ZD}
		B # Z # D	195	1					2		
		Residual	312	1							
Total	672										

there are two Residual or error terms, and hence the use in the past of split-unit-in-time. Here, the block-treatment interactions B # Z and B # Z # D can be tested. However, because Days are not randomized, a randomization test (Hinkelmann and Kempthorne, 2008, Section 6.5) is not available for Z # D, an interaction of interest (cf. the example in Section 4.1), and so any test must either use an assumed model or be a permutation test. This is not obvious from the split-unit-in-time analysis.

4.3 A repeated-measurement experiment with systematic allocation of treatments

Farewell and Herzberg (2003) describe a pain-rating experiment that is a two-phase experiment. Here, only the first phase, in which patients self-assess for pain while moving a painful shoulder joint, is considered. It involved eight patients, each of whom were observed on two occasions. Four of the patients had been judged to be expressive patients, having high levels of pain, and the other four had been judged to be unexpressive patients, having low levels of pain. Two motions are to be allocated to the two occasions for each patient: active motion performed by the patient without assistance and passive motion in which a therapist guided the patient's limb through its range of movement. The response is a pain rating made by each patient on each of the occasions that they underwent a shoulder motion. The experiment is a repeated-measurement experiment. It differs from the longitudinal example in Section 4.2 in having treatments (Motions) assigned to the time factor (Occasions) so that interest focusses on treatments, rather than times, differences. Because the design is already known, it is evaluated without formally employing the paradigm.

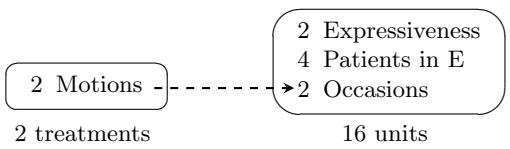

FIGURE 7 Factor-allocation diagram showing the treatments allocation to units for the Farewell and Herzberg pain experiment; the dashed arrow indicates that the allocation of Motions to Occasions is systematic; E = Expressiveness.

The factors for the experiment are Expressiveness, Patients, Occasions and Motions, with Motions allocated to the Occasions for each Patient. The factor Expressiveness is an inherent characteristic of each patient and so Patients is nested within Expressiveness. The factor Occasions is crossed with Expressiveness and Patients because, for all eight patients, the first occasion has the consistent property that it occurs before the second occasion. However, for the example, the motions were completed in the standard clinical order of active then passive (Brien, 2022). That is, they are systematically allocated

to Occasions and hence the dashed arrow in the factor-allocation diagram given in Figure 7.

The homogeneous allocation model is:

$$\text{Mean} + \text{M} + \text{E} + \text{M:E} + \text{O} + \text{O:E} \quad | \quad \text{Mean} + \text{E:P} + \underline{\text{E:P:O}}.$$

This model includes the term M:E, a term that gives rise to a block-treatment interaction. It also assumes that the units terms E and O:E are fixed.

TABLE 7 Anatomy tables for the homogeneous allocation models with Occasions (a) crossed with and (b) nested within Patients and Expressiveness for the Farewell and Herzberg pain experiment: E = Expressiveness; P = Patients; O = Occasions; M = Motions; EMS = Expected Mean Square; DF = Degrees of Freedom.

(a) Occasions crossed

units		treatments -blocks		EMS			
Source	**DF**	**Source**	**DF**	σ^2_{EPO}	σ^2_{EP}	σ^2_{μ}	$\theta.$
Mean	1	Mean	1	1	2	16	θ_{μ}
Expressiveness	1			1	2		θ_{E}
Patients [E]	6			1	2		
Occasions	1	Motions	1	1			$\theta^{\dagger}_{\text{O}\leftarrow\text{M}}$
O # E	1	M # E	1	1			$\theta^{\dagger}_{\text{OE}\leftarrow\text{ME}}$
O # P [E]	6			1			
Total	16						

(b) Occasions nested

units		treatments -blocks		EMS			
Source	**DF**	**Source**	**DF**	σ^2_{EPO}	σ^2_{EP}	σ^2_{μ}	$\theta.$
Mean	1	Mean	1	1	2	16	θ_{μ}
Expressiveness	1			1	2		θ_{E}
Patients [E]	6			1	2		
O [E:P]	8	Motions	1	1			$\theta^{\ddagger}_{\text{M}}$
		M # E	1	1			$\theta^{\ddagger}_{\text{ME}}$
		Residual	6	1			
Total	16						

†The subscripts of the θs indicate that the fixed source on the right of the left arrow is confounded with the fixed source on the left.
‡The sources for these θs are confounded with the source O [E:P] that has been omitted from the θ subscripts because both sources are only confounded with the random source O [E:P] and so neither θ involves O [E:P].

The anatomy table is in Table 7a. It shows that a test between patients with different Expressiveness is possible. However, Motions and Occasions are inextricably confounded with each other, *inextricable confounding* being a special form of exhaustive confounding in which there is only one allocated source confounded with the units source. The result is that their effects are inextricable in that they cannot be separately estimated. Further contributing to this is that both terms are fixed, as indicated by $\theta_{\text{O}\leftarrow\text{M}}$. Similarly, the two fixed terms M # E and O # E are inextricably confounded. Pseudoreplication of Motions is the cause, each Motion occurring on only one Occasion for each patient.

The inextricable confounding would be rectified if it is appropriate to assume that there is no Occasions main effect so that Occasions is nested within Patients. Along these lines, one might regard any order effect related to Occasions as being necessarily absorbed into the clinical procedure. The anatomy table for nested Occasions is in Table 7b. This analysis is the same as for a GRBD, with Expressiveness acting as a block factor and M # E being a block-treatment interaction of interest. Tests for both M and M # E are available in this analysis. Clearly, in designing this experiment a crucial question is whether the design needs to account for an Occasions order effect. The use of the function `designAnatomy` from the `R` package `dae` to produce the anatomy tables for the homogeneous allocation models, for both crossed and nested Occasions, is demonstrated in the `R` script in Supplement Section C.6. The EMSs were derived using the rules in the Appendix.

5 A nonorthogonal design

A sensory evaluation experiment is planned to assess eight food products using seven tasters, each of whom will evaluate all eight products.

The anticipated model. The factors involved in the experiment are Tasters, Evaluations and Products. It is thought that overall differences between both the Tasters and Evaluations are likely and that Tasters' evaluations of the Products may differ. Carry-over effects from a Taster's earlier evaluations

are not expected. Thus, the anticipated model is:

$$\text{Mean} + \text{P} \quad | \quad \text{Mean} + \text{T} + \text{E} + \underline{\text{T:P}} + \underline{\text{T:E}}. \tag{15}$$

Design selection & construction. Given the anticipated model, a row-column design is required and the Youden square design (YSD) for eight treatments (Cochran and Cox, 1957, Table 13.2) is the universally optimal design under the anticipated model or any alternative designation of Evaluations and Tasters as either fixed or random (Shah and Sinha, 1989). It is a *nonorthogonal design*, because it has canonical efficiency factors not equal to one. The appropriate permutations for a YSD are the same as for the LSqD. An `R` script to produce and plot a randomized layout is in Supplement Section C.7 and the plotted layout is in Supplement Figure 7.

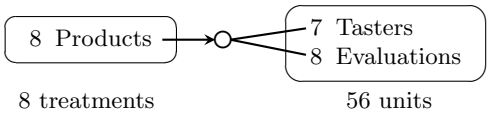

FIGURE 8 Factor-allocation diagram showing the treatments allocation to units for a Youden square design: the arrow indicates that the allocation is randomized; the '○' at the end of the arrow indicates that a nonorthogonal design was used; the two lines from the '○' indicate that the Products are allocated to the combinations of Tasters and Evaluations using the design. (cf. Figure 2)

The factor-allocation diagram. Figure 8 exhibits the factor allocation for the YSD.

The initial allocation model. The initial allocation model for this example is:

$$\text{Mean} + \text{P} \quad | \quad \text{Mean} + \text{T} + \text{E} + \underline{\text{T:E}}. \tag{16}$$

The homogeneous allocation model. The initial allocation model does not contain the term T:P for the block-treatment interaction that is in the anticipated model. Also, given that the A-efficiency for the design is 0.98, only 2% of the Products information is confounded with Evaluations; this information could be ignored in estimating the Products main effects, as would happen if Evaluations is assumed to be fixed. Thus, the homogeneous allocation model, based on the initial allocation model in (16), is:

$$\text{Mean} + \text{E} + \text{P} \quad | \quad \text{Mean} + \text{T} + \underline{\text{T:P}} + \underline{\text{T:E}}. \tag{17}$$

TABLE 8 Anatomy and traditional, skeleton ANOVA tables for the homogeneous allocation model for a YSD with a block-treatment interaction: T = Tasters; E = Evaluations; P = Products; EMS = Expected Mean Square; DF = Degrees of Freedom; Eff = A-Efficiency

(a) Anatomy table.

units		treatments-blocks			EMS				
Source	DF	Eff† Source	DF	σ^2_{TE}	σ^2_{T}	σ^2_{μ}	σ^2_{TP}	$\theta^{\ddagger}_{\cdot}$
Mean	1	1.00 Mean	1	1	8	56	1	θ_{μ}
Tasters	6			1	8		1	
Evaluations	7	0.02 $\text{Products}_{\text{E}}$	7	1			1	$\theta_{\text{E}\leftarrow\text{P}}$
T # E	42	0.98 $\text{Products}_{\text{TE}}$	7	1			1	$\theta_{\text{TE}^*\leftarrow\text{P}}$
		1.00 T # P	35	1			1	
Total	56							

(b) Traditional, skeleton ANOVA table.

		EMS			
Source	DF	σ^2_{TP}	σ^2_{T}	σ^2_{μ}	$\theta^{\ddagger}_{\cdot}$
Mean	1	1	8	56	θ_{μ}
Tasters	6	1	8		
Evaluations	7	1			θ_{E}
Products (Adj.)	7	1			$\theta_{\text{P(Adj)}}$
T # P	35	1			
Total	56				

†Each Eff is the A-efficiency value for a treatments source when it is confounded with the units source aligned with or immediately to the left.

‡Each $\theta_{\cdot}$ is a quadratic form in the expectation of a response variable, with the subscript indicating the source whose scaled projection matrix is the matrix of the quadratic form. The subscript $\text{E}\leftarrow\text{P}$ connotes that the matrix of the quadratic form is the scaled projection matrix for Products confounded with Evaluations and, as both source names do not have asterisks, they are fixed sources. On the other hand, $\theta_{\text{TE}^*\leftarrow\text{P}}$ indicates the matrix of the quadratic form is the scaled projection matrix for Products confounded with T # E and that T # E is a random source, so that only Products will contribute to the expectation.

Anatomy table. The anatomy table is in Table 8a. It includes a major column that is labelled treatments-blocks because it involves the block-treatment interaction T:P. Being a nonorthogonal design, the table involves *partial confounding* that is exposed in the table by a treatments source being confounded with multiple units sources. To portray how much treatment information is associated with each units source, the table includes a column of A-efficiency values for treatments-blocks sources. All of the other designs in this paper have their A-efficiency column omitted because their A-efficiency values are all one, i.e. they are orthogonal.

Table 8a shows that, while no Products information is confounded with Tasters, Products is partially confounded with both Evaluations and T#E, there being two Products sources that are differentiated by subscripts that are the abbreviated names for the sources with which they are partially confounded; the amounts of Products information confounded with E and T # E are 2% and 98%. Each of the Products sources has seven canonical efficiency factors and so their DFs are seven. The number of unique values of the canonical efficiency factors for a treatments(-blocks) source confounded with a units source is called its *order of balance*. For *balanced designs*, such as the YSD, all treatments(-blocks) sources are order one and the single value of the canonical efficiency factors for a treatments(-blocks) source is its A-efficiency. The smallest entity with which Products is confounded is T # E, each entity being a single unit. An A-optimal design must have the maximum possible A-efficiency for $\text{Products}_{\text{TE}}$; the YSD being A-optimal, 0.98 is the maximum possible for a design of 7 rows $\times$ 8 columns for 8 treatments.

There is also partial aliasing in that only 35 of the 42 DF for T # P is estimable, the other seven DF being aliased with Products when estimated from the source T # E. It is termed aliasing, rather than confounding, because it involves two sources, $\text{Products}_{\text{TE}}$ and T # P, that are associated with the same panel; both sources include Products. However, the estimable part of T # P is confounded with T # E.

Because of the partial aliasing, the EMS rules in the Appendix cannot be used directly. However, it is shown in Supplement Section F that the rules can be applied by transferring σ^2_{TP} from treatment to units, and adding it to σ^2_{TE} (see also Example 3 in the Appendix). The EMSs in Table 8a reveal that (i) σ^2_{TE} and σ^2_{TP} are not separately estimable, (ii) the fixed effects of E and $\text{Products}_{\text{E}}$ cannot be separately estimated and so the Products main effect must be estimated using $\text{Products}_{\text{TE}}$ which is of little consequence seeing that $\text{Products}_{\text{TE}}$ contains 98% of the Products information, and (iii) an F-test for the Products main effect would use the ratio of the $\text{Products}_{\text{TE}}$ mean square to the T # P mean square. The use of the function `designAnatomy` from the `R` package `dae` to produce the anatomy table, with A-efficiency criteria and the orders of balance, but without EMSs, is demonstrated in the `R` script in Supplement Section C.7. Also produced is a table that details the partial aliasing.

Table 8b shows the traditional, skeleton ANOVA. This ANOVA decomposition is equivalent to that in the anatomy table. In particular, the Products (Adj.) source and the $\text{Products}_{\text{TE}}$ source are the same and $\theta_{\text{P(Adj)}}$ equals $\theta_{\text{P}\leftarrow\text{TE}^*}$. However, Table 8b does not provide the insights into the properties of the experiment to same extent as Table 8a: it does not reflect the confounding, draw attention to the partial aliasing of T # P or separate inherent random variability and block-treatment interaction, the terms T # E and T # P being merged into a single source. Indeed, the terms in Table 8b are fitted in order from top to bottom and all intermingling of information may well be characterized as aliasing, rather than confounding (see for example Chambers et al., 1992, Section 5.3.2).

For this example, the homogeneous allocation model is singular and so, to be able to fit a model, a nonsingular model must be identified. Omitting either T:E or T:P will result in a nonsingular model. The removal of T:P results in the following nonsingular model:

$$\text{Mean} + \text{E} + \text{P} \quad | \quad \text{Mean} + \text{T} + \underline{\text{T:E}}.$$

For this model, the estimate of σ^2_{TE} actually estimates $\sigma^2_{\text{TE}} + \sigma^2_{\text{TP}}$.

6 Discussion

The paradigm outlined in Figure 1 provides a conceptual framework for designing experiments. While we acknowledge that strict adherence to the steps outlined is not always necessary, the use of its core elements, the identification of the treatments and units factors, the inclusion of block-treatment interactions and the production of an anatomy table with EMSs, provides a designer with the following benefits that are not available using the traditional ANOVA approach:

- An appreciation of the distribution of treatments information across the units sources, this being provided by an explicit depiction of the confounding in a design. As well as gaining an understanding of the design's properties, it aids in the evaluation of its suitability for an experiment.
- A capacity to check that the randomized layout for an experiment has been correctly generated. Typos and programming mistakes can be detected.
- The ability to separately consider the likelihood of block-treatment interactions and sources of inherent random unit variability at the design stage, and, if necessary, allow for them in the design.
- An automatic warning in the anatomy table when the design involves pseudoreplication.
- The means to display singularities in the variance model.

To achieve these benefits does require keeping to the discipline of including ALL units factors, rather than just blocking factors, in the description and models of a design. It is the authors' experience that a failure to do so often results in mistakes in the description of the experiment and the analysis. For instance, in Section 4.3, Occasions has not been included in previous descriptions of this experiment and so the systematic allocation and the pseudoreplication of Motions had not been recognized. At the very least, the pseudoreplication should be acknowledged when reporting the experiment.

The major issues highlighted by the paradigm are now summarized.

Anticipated effects determine the design for an experiment. Whether or not an anticipated model is explicitly formulated, the anticipated effects are crucial to the choice of a design for an experiment. The factor allocation involved in the chosen design will need to be restricted to reflect the crossing and nesting of factors inherent in the anticipated effects, rather than the crossing and nesting of the factors intrinsic to the physical set-up of an experiment. That is, the nesting and crossing of factors arising from the allocations should match that associated with anticipated effects, as has been demonstrated using alternative designs for a 5×5 grid of plots in Section 2.

Understanding the confounding intrinsic to a design is valuable. Using software, such as the `R` package `dae`, to produce an anatomy table that exhibits the confounding arising from a proposed layout for an experiment is crucial to understanding the confounding. We urge designers to always produce the anatomy table for a proposed design, even if other aspects of the paradigm are not adopted, and to do this using software, rather than manually, in order to validate one's perception of the confounding. It is particularly appropriate for designs whose models have multiple errors, such as designs with split-units, repeated measurements designs and nonorthogonal designs.

Fundamentally, the confounding indicates how the treatment information is distributed between the units sources, from which the proper error for each treatments source can be gleaned. This is equivalent to identifying the experimental unit, because an *experimental unit* for a treatment factor can be defined as the smallest entity, identified by the units factors in a source, with which a treatment source is confounded. For example, the experimental unit for Zinc for the GRBD in Section 3.2 is a pot in a block, Zinc being confounded with Pots [B]. However, this requires that the factors are separated into treatments and units factors and ALL units factors for the design are included. The alternative where some units factors are omitted so that experimental units must be specified using block-treatment interactions is problematic: (i) it is incongruous to speak of allocating treatments to block-treatment interactions; (ii) the experimental units do not always correspond to block-treatment interactions, for example, for the GRBD. Further, it leads to all variation between experimental units being interpreted as block-treatment interactions and obscures the contribution of inherent unit variability.

A particular advantage arises when a design involves pseudoreplication so that a treatments effect is inextricably confounded with a units effect (e.g. Section 4.3). This characterization of the problem with pseudoreplication is more understandable to researchers than the more usual explanation of the treatments not being appropriately replicated; it directly addresses the problem that arises.

Block-treatment interaction differs from inherent random unit variability. The paradigm distinguishes between these two different types of variation and allows different strategies to be employed with respect to the inclusion of block-treatment interactions in linear mixed models for comparative experiments. Two extreme strategies are: (i) always assume block-treatment additivity and (ii) always include all block-treatment interactions. A motivation for employing (i) is that, when all allocations are randomizations, it produces a randomization model and randomization tests are available. Strategy (ii) is perhaps driven by the practice of identifying experimental units in terms of block-treatment interactions. Here we have employed an intermediate strategy of only including block-treatment interactions that are thought to be relevant for the current experiment, this being consistent with the strategy of formulating an anticipated model. They are particularly relevant when units factors are of interest to

the researcher, as for the examples in Sections 4.2 and 5. Nonetheless, the paradigm that we have outlined has the flexibility to allow one to choose any of the three strategies as a default strategy.

It is left to the reader to choose their own strategy, but the inclusion of block-treatment interaction can affect the properties of a design. Three situations arise: (i) no differences between analyses based on additivity and on nonadditivity (Section 3.1); (ii) only some block-treatment interactions can be estimated separately (Section 4.1); (iii) block-treatment interactions can be estimated separately from inherent random variability (Section 3.2). For (ii), it does not seem satisfactory to test only some of the block-treatment interactions. More serious effects from the addition of block-treatment interactions are: (i) hypothesis tests for effects of interest may entail denominator mean squares with reduced DF, and (ii) randomization tests are no longer appropriate.

Allocation differences elucidate experiment differences. The three experiments in Section 4 have in common that time is a factor. However, they differ in whether the time factor is an allocated or a recipient factor and whether treatment factors are allocated to the time factor. Only when the time factor was allocated, as in Section 4.1, was a split-unit design used and randomization inferences involving the time factor are possible. The other two experiments are repeated-measurements experiments, in which observations are made on an entity over time and/or space, and the time factor is a recipient factor. The example in Section 4.2 is a longitudinal experiment because there are no treatments allocated to the time factor and, as in the split-unit experiment, the time differences are of interest. For the example in Section 4.3, the time factor Occasions is of no interest.

Factor allocation provides the basis for determining a linear mixed model. The terms in a linear mixed model should be based on the factor allocation used for the design to ensure that the terms appropriate to this allocation are included in the initial allocation model. The first benefit of this is that, if all allocations are randomized, the initial allocation model will be equivalent to a randomization model. Secondly, an anatomy based on the homogeneous model will reflect the confounding induced by the allocations, the only difference between the initial and homogeneous allocation model, as far as what terms they contain, being that the latter model includes block-treatment interactions. These interactions are incorporated to account for effects that might occur in practice. Also, as Brien and Demétrio (2009) suggest, a third stage of model modification may be necessary to produce an initial model for data analysis that they term a *prior allocation model*; in this stage, model terms may be reparametrized to include trend terms and/or random effects that are not assumed to be independent and identically distributed. The longitudinal experiment in Section 4.2 is an example where this is likely to be the case.

Pedagogical advantages. We assert that, given the benefits outlined above, our approach to designing experiments has pedagogical advantages. In particular, with the ongoing shift from ANOVA to mixed model fitting as the prime method of analyzing an experiment's data, an understanding of how treatments information is distributed across the units sources for a design has become increasingly difficult to acquire. We argue that assessing designs using anatomy tables can help to reverse this trend.

Support in R. The paradigm is supported by the `R` package `dae` (Brien, 2023), as is illustrated in Supplement Section C. The functions in `dae` require a systematic design to be supplied, which is straightforward for orthogonal designs like RCBDs, LSqDs and split-unit designs, including the Loughin (2005) design. For nonorthogonal designs either a catalogue of designs or software that searches for optimal designs (`odw` (Butler, 2022) is an `R` package) is required. In the case of computer searching for optimal designs, it is commonly thought that, because random interchanges of treatments are employed, the design has been randomized. However, while a random design may have been generated, it cannot be guaranteed that is has the property required for it be properly randomized, namely that, given the restrictions of the blocking structure, all treatments had the same probability of being allocated to each unit. Their proper randomization can be ensured by using the function `designRandomize` from `dae`.

For more complicated design, it can be useful to verify that the appropriate model terms have been specified, given the nesting and crossing of factors inherent in the anticipated model and which is captured in the factor allocation diagram; this has been achieved in the `R` scripts in Supplement Section C by using crossing and nesting operators in the `designAnatomy` functions calls.

References

Addelman, S. (1969) The generalized randomized block design. *The American Statistician*, *23*, 35–36.

Bate, S. T. and M. J. Chatfield (2016a) Identifying the structure of the experimental design. *Journal of Quality Technology*, *48*, 343–364.

Bate, S. T. and M. J. Chatfield (2016b) Using the structure of the experimental design and the randomization to construct a mixed model. *Journal of Quality Technology*, *48*, 365–387.

Brien, C. J. (1983) Analysis of variance tables based on experimental structure. *Biometrics*, *39*, 53–59.

Brien, C. J. (1989) A model comparison approach to linear models. *Utilitas Mathematica*, *36*, 225–254.

Brien, C. J. (2017) Multiphase experiments in practice: A look back. *Australian & New Zealand Journal of Statistics*, *59*, 327–352.

Brien, C. J. (2019) Multiphase experiments with at least one later laboratory phase. II. Northogonal designs. *Australian & New Zealand Journal of Statistics*, *61*, 234–268.

Brien, C. J. (2022) Designing, understanding and modelling two-phase experiments with human subjects. *Statistical Methods in Medical Research*, *31*, 626–645.

Brien, C. J. (2023) *dae: functions useful in the design and ANOVA of experiments*. Available at https://CRAN.R-project.org/package=dae/, (R package version 3.2-14).

Brien, C. J. and C. G. B. Demétrio (2009) Formulating mixed models for experiments, including longitudinal experiments. *The Journal of Agricultural, Biological, and Environmental Statistics*, *14*, 253–280.

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. *The Journal of Agricultural, Biological, and Environmental Statistics*, *16*, 422–450.

Butler, D. G. (2022) *odw: Generate optimal experimental designs*. Available at https://mmade.org/, (R package version 2.1.4).

Chambers, J. M., A. Freeny, and R. M. Heiberger (1992) Analysis of variance; designed experiments. In J. M. Chambers and T. J. Hastie (Eds.), *Statistical Models in S*. Chapter 5, pp. 145–193. Pacific Grove, California: Wadsworth & Brooks.

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

Farewell, V. T. and A. M. Herzberg (2003) Plaid designs for the evaluation of training for medical practitioners. *Journal of Applied Statistics*, *30*, 957–965.

Finney, D. J. (1943) The fractional replication of factorial arrangements. *Annals of Eugenics*, *12*, 291–301.

Fisher, R. A. (1935a) Contribution to the discussion of "Complex experiments" by F. Yates. *Supplement to the Journal of the Royal Statistical Society*, *2*, 229–231. Incomplete extract at https://hdl.handle.net/2440/15225/.

Fisher, R. A. (1935b) *The Design of Experiments*. (1st ed.) Edinburgh: Oliver and Boyd.

Harville, D. A. (1991) Contribution to the discussion of a paper by M.L. Samuels, G. Casella and G.P. McCabe. *Journal of the American Statistical Association*, *86*, 812–815.

Hinkelmann, K. and O. Kempthorne (2008) *Design and Analysis of Experiments: Volume I: Introduction to Experimental Design*. (2nd ed.) New York: John Wiley & Sons.

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.

Littell, R. C., G. A. Milliken, W. W. Stroup, R. D. Wolfinger, and O. Schabenberger (2006) *SAS for Mixed Models*. (2nd ed.) Cary: SAS Press.

Loughin, T. M. (2005) Seven error terms!! are you kidding?? Paper presented at JSM 2005, Minneapolis, USA.

Mead, R., S. G. Gilmour, and A. Mead (2012) *Statistical principles for the design of experiments*. Cambridge: Cambridge University Press.

Milliken, G. A. and D. E. Johnson (1992) *Analysis of Messy Data I. Designed Experiments*. New York: Chapman & Hall/CRC.

Montgomery, D. C. (2013) *Design and Analysis of Experiments*. (8th ed.) Hoboken, NJ: John Wiley & Sons Inc.

Nelder, J. A. (1965a) The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. *Proceedings of the Royal Society, Series A*, *283*, 147–161.

Nelder, J. A. (1965b) The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. *Proceedings of the Royal Society, Series A*, *283*, 162–178.

Piepho, H. P., A. Büchse, and K. Emrich (2003) A hitchhiker's guide to mixed models for randomized experiments. *Journal of Agronomy and Crop Science*, *189*, 310–322.

Piepho, H. P., A. Büchse, and C. Richter (2004) A mixed modelling approach for randomized experiments with repeated measures. *Journal of Agronomy and Crop Science*, *190*, 230–247.

R Core Team (2023) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.r-project.org/.

Samuels, M. L., G. Casella, and G. P. McCabe (1991) Interpreting blocks and random factors. *Journal of the American Statistical Association*, *86*, 798–821.

SAS Institute Inc. (1989–2022) *JMP®, version 17.0*. Cary, NC: SAS Institute Inc. Available at https://www.jmp.com/en_us/software/data-analysis-software.html.

SAS Institute Inc. (2020) *SAS/QC® 15.2 User's Guide*. Cary, NC: SAS Institute Inc. Available at https://support.sas.com/en/software/sas-qc-support.html#documentation/.

Shah, K. R. and B. K. Sinha (1989) *Theory of optimal designs*. New York: Springer-Verlag.

Steel, R. G. D. and J. H. Torrie (1960) *Principles and Procedures of Statistics*. New York: McGraw-Hill.

Stroup, W. W. (2012) *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications*. Boca Raton: Chapman & Hall/CRC.

VSN International (2022a) *CycDesigN for Windows, version 8.0.* Hemel Hempstead, UK: VSN International. Available at http://www.vsni.co.uk/software/cycdesign/.

VSN International (2022b) *Genstat 22nd Edition.* Hemel Hempstead, UK: VSN International. Web page: http://www.genstat.co.uk/.

Wilk, M. B. and O. Kempthorne (1956) Some aspects of the analysis of factorial experiments in a completely randomized design. *Annals of Mathematical Statistics*, *27*, 950–985.

Acknowledgements

The authors would like to thank the associate editor and reviewers whose comments and questions helped to improve the readability and clarity of the paper.

Conflict of Interest

The authors have declared no conflict of interest.

Data availability statement

The `R` scripts for all of the examples presented in the paper are presented in Supplement Section C and are available, along with their output, in the supplementary zip file.

Appendix Rules for obtaining the expected mean squares for a design using an anatomy table

The rules for obtaining the expected mean squares (EMSs) given here apply to experiments in which the anatomy table is formed from a set of units sources and a set of treatments sources for which (i) the sources within each set of sources are mutually orthogonal, (ii) the set of treatment sources is structure-balanced with respect to the joint decomposition of the sets of treatments and units sources (Brien, 2019), (iii) fixed and random terms are not constrained to correspond to treatments and units terms, respectively, and (iv) allow for when a fixed source is confounded with other fixed sources. The rules may apply when these conditions are not met. For example, when the design is not structure-balanced because it involves aliasing, but it is first-order balanced because all of the orders for the design are one (for an example see Section 5). They are also likely to apply when any unbalanced sources are for fixed treatment terms only. The derivation of these rules is outlined in Supplement Section E.

These conditions can be checked with the `designAnatomy` function in the `R` (R Core Team, 2023) package `dae` (Brien, 2023). For all structure-balanced designs, all `orders` are one and there is no partial aliasing; an orthogonal design is a structure-balanced design for which all `aefficiency` values produced by `dae` are one.

The anatomy tables that have been presented in this paper consist of three major columns, labelled as follows: (i) units, (ii) treatments or treatments-blocks, and (iii) EMS. In a row, there will be either (i) just a units source, (ii) the Residual of a units source or (iii) a units source and a treatments(-blocks) source that is confounded with the units source; the source for (iii) is a joint source. For each of row, there is a single EMS.

The rules are now presented and are followed by some examples of their application.

For *each row* in the anatomy table: The EMS consists of the sum of contributions for each source in the row, these sources being the right-most source in the row and those sources with which it is confounded, if any; that is, contributions from either a single units source or two sources that come from different major columns in the table. Determine the EMS for a row as follows:

Obtain the random contribution for each contributing source in the row: The *random contribution* for a source is a linear combination of the variance components for terms from the same major column as the source.

Beginning with the left-most column of sources and continuing across to the right-most source for that row, obtain the linear combination that is the random contribution for each of these sources (ignore Residual sources) as follows:

1. For the current source, identify its term: it is comprised of all factors in the current source and it is referred to as the current term. Both are represented by abbreviated names formed

from the initial capital letters of the factors that comprise them. If the source is not from the first major column and it is confounded with the units source for the current row, ascertain the value of the A-efficiency criterion for the current source in the current row, the value being the harmonic mean of its canonical efficiency factors.

2. For the current source, determine its random contribution to the EMS for the row: it is the linear combination of the variance components for all random terms to which the current term is marginal; i.e. the column space for the current term will be a subspace of the column space for each of these random terms.

   If the current term is random, also add its variance component to the linear combination.

   The coefficient of each variance component in the linear combination is the number of replicates of the observed combinations of the levels of the factors in the component's random term; it is calculated as the number of observations for which the experimental design has been generated divided by the number of observed combinations of the levels of the factors in the component's random term.

   If the current source is not a units source, then the linear combination is now multiplied by the value of the A-efficiency criterion for the current source in the current row.

Form the EMS *for the row*: It is the sum of the random contributions of its sources, i.e. the sum of the linear combinations, to which is added a fixed contribution, provided at least one fixed source is involved in the row.

The *fixed contribution for a row*, θ, is a quadratic form in the expectation for a response. Here, the θ for a row is specified by nominating the matrix of the quadratic form for the row using a subscript that is a list of the abbreviated names for sources in the row, beginning with the left-most source and proceeding to the right-most source and with sources separated by left arrows ($\leftarrow$). The left arrow indicates that the source at the tail is confounded with the source at the head of the arrow; thus, the matrix of the quadratic form is the projection matrix for the part of the source at the head that pertains to the source at the tail. All sources are fixed unless they are asterisked, in which case they are random sources; note that while the projection matrix may incorporate random sources, only fixed sources have terms that contribute to the expectation.

Suppose that we have allocated source A confounded with recipient source B. Thus, $\theta_{\mathrm{B}^*\leftarrow\mathrm{A}}$ indicates that the matrix of the quadratic form is the scaled projection matrix for the part of B that pertains to A, where A is fixed and B is random. Consider the θs for the four combinations of the A and B sources being either fixed or random:

Both fixed: $\theta_{\mathrm{B}\leftarrow\mathrm{A}}$;
Only A fixed: $\theta_{\mathrm{B}^*\leftarrow\mathrm{A}}$ or, if A is only confounded with B, θ_{A}; while B may affect the projection matrix for the quadratic form, it is not involved in the expectation;
Only B fixed: $\theta_{\mathrm{B}\leftarrow\mathrm{A}^*}$ and $\theta_{\mathrm{B}_{\vdash}}$, where $\mathrm{B}_{\vdash}$ means that part of the recipient source B that is orthogonal to all allocated sources; A only plays a role in the projection matrix for the quadratic form, it making no contribution to the expectation;
Both random: no θ is applicable.

If there is no allocated source confounded with B and B is fixed, then the θ is θ_{B}.

Example 1: Take the Rows line in the body of Table 1a; its right-most source is the Rows source and it is the only source.

1. We begin with the left-most source Rows that comes from the units major column. Its term is R and an A-efficiency criterion does not apply. The only random term in the units major column to which R is marginal is R:C. Also, R is a random term. Then the random contribution for the current source is a linear combination of the variance components for R:C and R. The numbers of replicates for these terms are 1 and 5. Thus, the contribution of Rows is $\sigma^2_{\mathrm{RC}} + 5\sigma^2_{\mathrm{R}}$.
2. There are no other sources in this line to contribute. The EMS for the line is just the Rows contribution, as shown in Table 1a.

Example 2: Take the line in Table 1a that includes Varieties. Its right-most source is Varieties and is confounded with R # C.

1. We begin with the left-most source R # C, that is a units source. Its term is R:C and an A-efficiency criterion does not apply. It is a random units term and, being an identity term, there

is no random term to which it is marginal; so it is the only random contribution for this source. The number of observed combinations of the factors in R:C equals the number of observations. Thus, the coefficient of its variance component is 1 and the contribution of R # C is σ^2_{RC}.

2. The next source in the line is Varieties from the treatments major column. Its term is V and it is a fixed treatments term. There are no random terms from the treatments major column to which it is marginal, it being the only treatments term. Thus, there is no random contribution from the treatments major column in this line of the table.
3. The only fixed source in the current line of the table is Varieties and its contribution is $\theta_{\mathrm{RC}^*\leftarrow\mathrm{V}}$; given that Varieties is only confounded with R # C, the fixed contribution simplifies to θ_{V}.
4. The EMS for the line is the sum of the random and fixed contributions, as shown in Table 1a;

Example 3: Take the row of Table 8a that contains the source $\mathrm{Products_E}$. It is the right-most source and is confounded with Evaluations. The rules apply, even though T # P is partially aliased with other treatments sources, because if this source is omitted then the design is structure-balanced.

1. We begin with the left-most source Evaluations, which is a units source. Its term is E and an *A*-efficiency criterion does not apply; it is a fixed term. The only units term to which it is marginal is T:E. Thus the random contribution for the current source is the variance component for T:E, with coefficient one.
2. The next source is $\mathrm{Products_E}$, a treatments source, whose term is P. It is a fixed term and the source has *A*-efficiency 0.02. The only random treatments term to which it is marginal is the identity term T:P. Thus, the random contribution for the current source is the variance component for T:P, with coefficient one.
3. The combined random contribution to the EMS is $\sigma^2_{\mathrm{TE}} + \sigma^2_{\mathrm{TP}}$.
4. There is a units and a treatments fixed source involved in this row of Table 8a. The fixed contribution θ for this row is $\theta_{\mathrm{E}\leftarrow\mathrm{P}}$. Thus, the EMS for the row is the sum of the combined random contribution and the fixed contribution: $\sigma^2_{\mathrm{TE}} + \sigma^2_{\mathrm{TP}} + \theta_{\mathrm{E}\leftarrow\mathrm{P}}$, as shown in Table 8a.