Multiple randomizations

C. J. Brien†

University of South Australia, Adelaide, Australia.

R. A. Bailey

Queen Mary, University of London, London, United Kingdom.

[Received February 2, 2005. Final revision December 9, 2005]

Summary. Multitiered experiments are characterised by involving multiple randomizations, in a sense that we make explicit. We compare and contrast six types of multiple randomizations, using a wide range of examples, and discuss their use in designing experiments. We outline a system of describing the randomizations in terms of sets of objects, their associated tiers and the factor nesting, using randomization diagrams, which give a convenient and readily assimilated summary of an experiment's randomization. We also indicate how to formulate a randomization-based mixed model for the analysis of data from such experiments.

Keywords: Design of experiments; Mixed models; Multiple randomizations; Multitiered experiments; Pseudofactors; Randomization; Superimposed experiments; Tiers; Two-phase experiments

1. Introduction

Experiments are distinguished from observational studies and happenstance data by the purposive application of treatments to observational units. Nelder (1965a,b), White (1975), Bailey (1981, 1991) and Heiberger (1989) formulated methods for the analysis of experiments that take this distinction into account by classifying factors in the experiment as either 'block' or 'treatment' factors. Here we interpret 'block' and 'treatment' factors to mean the sets of unrandomized and randomized factors, respectively. Many authors have advocated the use of such a distinction (Fisher, 1935; Wilk and Kempthorne, 1957; Cox, 1958, Section 6.3; Yates, 1975; Mead and Curnow, 1983, Section 14.1; Brien, 1983; Piepho, Büchse and Emrich, 2003). The distinction enables the direct construction of analysis-of-variance tables exhibiting all the confounding, as well as the inclusion, in a mixed model, of all terms warranted by the randomization.

Of course, it is possible to formulate an analysis without making this distinction. For example, in the analysis of a randomized complete block design as a 'two-way anova' based on a two-factor, no-interaction model, no distinction is made between the treatments that are randomly assigned and the blocks that result from inherent features of the observational units. As Kempthorne (1955) notes, this results in the same analysis for the randomized complete block design and the two-way factorial experiment, even though they differ markedly in the randomization employed. This is because the confounding inherent in each experimental situation has not been recognized (Brien and Payne, 1999). Similarly,

†*Address for correspondence:* Dr C.J. Brien, School of Mathematics and Statistics, University of South Australia, North Terrace, Adelaide, SA 5000, Australia
E-mail: chris.brien@unisa.edu.au

the common description of the classic split-plot experiment in terms of only the three factors blocks, main-plot treatments and subplot treatments also fails to distinguish between randomized treatments and inherent features of the observational units.

Brien (1983) identified experiments that involve more than the single randomization of 'treatments' to 'plots' and concluded that two sets of factors are not enough to describe the randomization in such experiments. The general term 'tiers' was introduced for the sets of factors in an experiment which result from the classification of the factors according to their status in the randomization: see also Cullis et al. (2003). Experiments that involve multiple randomizations, and hence more than two tiers, were labelled 'multitiered'. Multitiered experiments include two-phase, some superimposed and some single-stage experiments, and some multistage experiments using the same units at each stage; they do not represent a collection of new designs, but are a class of designs made up of several existing design types.

Two-phase experiments were introduced by McIntyre (1955) and were discussed by Cox (1958), although Cox used the term 'stage' rather than 'phase'. They are characterized by

(a) a complete experiment in the first phase, although not necessarily with the measurement of response variables;
(b) the randomization of the units from the first phase to the units in the second phase.

A common situation in which they occur is where the produce from a field trial has to be processed in a laboratory or an evaluation phase (Brien, 1983; Brien, May and Mayo, 1987; Wood, Williams and Speed, 1988; Brien and Payne, 1999; Cullis et al., 2003; Kerr, 2003). All of the examples of two-phase experiments published in the literature employ only one of the simplest type of multiple randomization.

In this paper we distinguish between stages and phases in experiments, the former including the latter as a special type. We note that usage of stages and phases in experiments is unrelated to that in sampling. Multistage experiments are conducted in several distinct time intervals with randomization of factors for each interval. They include two-phase, superimposed and changeover experiments (but not longitudinal studies), as well as multistage reprocessing experiments, in which the response is measured only after several stages of processing the same units (Miller, 1997; Mee and Bates, 1998; Box and Jones, 1992, Section 5). Not all multistage experiments are multitiered as they do not involve multiple randomizations—the usual repeated measurements experiments discussed in textbooks are not. Single-stage experiments that involve multiple randomizations include some grazing and some plant experiments. For example, Brien and Demétrio (1998) discuss multiple randomizations in the context of continuous grazing trials and Preece (1991) describes a multitiered horticultural experiment.

Although the experiments do not involve new designs, there has been uncertainty and difficulty in their randomization and analysis. The purpose of this paper is to describe and compare six different types of multiple randomization and how they are employed in a range of multitiered experiments. We concentrate on equi-replicate designs, because they demonstrate all the relevant issues but without extra complications. We hope to increase understanding of the way in which multiple randomizations can be employed in designing experiments. We also develop randomization diagrams for depicting the randomization in an easily absorbed form. In Section 2, we introduce and define the terms required and illustrate their use in the context of two-tiered experiments. Section 3 describes a simple three-tiered experiment. The six types of multiple randomizations are then described: Section 4 discusses and gives examples of multiple randomizations in which the two constituent randomizations can be performed in either order, while Section 5 does the same for the inclusive multiple randomizations. Section 6 gives examples involving more than one type of multiple randomizations, and hence three or more randomizations. A strategy for formulating a mixed model for analysing multitiered experiments is outlined in Section 7. Section 8 addresses a number of general issues concerning terminology and multiple randomizations. Further examples involving multiple randomizations are available in Brien (1992) and from the multitiered experiment web site at <`http://chris.brien.name/multitier/`>.

2. The elements of a single randomization

2.1. A simple example: first randomization

Example 1 (a two-phase sensory experiment): Brien (1983) describes a two-phase sensory experiment (McIntyre, 1955) to evaluate a set of wines made from the produce of a field trial in order to test the effects of several viticultural treatments. The first phase is a field trial in which the t treatments are to be assigned to vines arranged in b blocks of t plots of vines. How should such an experiment be randomized? Often the answer given is: in each block independently, choose a random order for the t treatments. We prefer to think of this as first making a systematic design with each treatment on one plot per block and secondly permuting the plot (labels) by randomly permuting the blocks and then randomly permuting the plots within each block independently. One advantage of the second approach is that the method of starting with a systematic design and then choosing a random permutation of the plots, possibly subject to some restrictions, generalizes to almost all single randomizations. A second advantage is that, under suitable conditions (Nelder, 1965a; Bailey, 1981, 1991), the group of allowable permutations of the plots defines a covariance matrix on the observations whose eigenspaces are independent of the values of the correlations and variances. These eigenspaces are called *strata*, and they define the structure of the analysis of variance.

In this example the permutation group is the wreath product of the symmetric group on blocks with the symmetric group on plots, written $S_t \wr S_b$ for b blocks with t plots each. The assumption of plot-treatment additivity and the method of randomization allow us to assume that the expectation of the response Y_α on plot α is a constant depending only on the treatment $f(\alpha)$ applied to that plot, and that the covariance between Y_α and Y_β takes only three values: σ^2 if $\alpha = \beta$, $\rho_1\sigma^2$ if α and β are different plots in the same block; and $\rho_2\sigma^2$ otherwise. This then specifies the randomization model on which a randomization analysis of the experiment can be based.

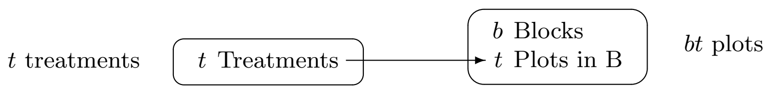

Fig. 1. Randomization diagram for a randomized complete block design

The randomization diagram in Fig. 1 succinctly summarizes the design and the appropriate randomization. On the left is a *panel* showing the set of treatments: there are simply t treatments, with no further structure. On the right is a panel for the set of bt plots, whose structure is defined by a b-level Block factor and a t-level Plot factor nested within Blocks. This structure on the plots defines the group $S_t \wr S_b$ used in the randomization. The arrow shows the systematic design: one treatment on each plot within each block.

2.2. Defining the concepts

This simple example enables us to introduce many of our concepts. A randomization involves the random assignment of one set of objects to another set of objects, e.g. a set of t treatments to a set of bt plots. In each case the name and number of the objects are written outside the panel in the randomization diagram, as in Fig. 1; the names are written in lower case letters to distinguish them from factor names, that have initial capitals.

Brien (1983) defined a *tier* to be a set of factors on one of the sets of objects in a randomization. Levels of the factors in different tiers connected by arrows must be associated by randomization whereas those in the same tier are not. In the simple example we have the two tiers {Treatments} on treatments and {Blocks, Plots} on plots—each tier simply consists of the factors in one of the panels in Fig. 1. In this example, the levels of Treatments are associated with particular levels of Plots in Blocks by randomization and so Treatments and Plots are in different tiers. The levels of Blocks are not associated with the levels of Plots in Blocks by randomization and they are in the same tier.

Each set of objects is uniquely indexed by all the levels-combinations of all the factors in the tier. Following Bailey (1996) and Tjur (1984), we use $F_1 \wedge \cdots \wedge F_n$ to denote what we term the *generalized factor* whose levels are the levels combinations of F_1, F_2, ... and F_n, for $n \geq 1$. If a generalized factor contains F_i and is to be meaningful, then it must also contain every F_j which nests F_i. We consider only such generalized factors, calling them *intrinsic*. In Example 1, Blocks $\wedge$ Plots is the bt-level generalized factor that uniquely indexes plots. Thus, if Blocks has level i and Plots has level j then Blocks $\wedge$ Plots has level (i, j). On the treatments tier, Treatments is the only intrinsic generalized factor, while on the plots tier, the intrinsic generalized factors are Blocks and Blocks $\wedge$ Plots.

There is a possible ambiguity about what a factor is. Does the plot factor in Example 1 have t levels or bt? The latter convention is more general; for example, it allows for blocks of unequal size. The former convention, in which the plot factor is called a *prefactor* by Bailey (1991), is more convenient for computers so long as nesting relationships are specified. In this paper we use the former convention, which is consistent with the usage of Nelder (1965a) and Heiberger (1989), as well as many textbooks and statistical computing packages.

To make explicit the nesting relationships between the factors within each tier, they too are shown in randomization diagrams, such as Fig. 1, by listing, with each factor, the list of factors that nest it. A tier with a single factor is drawn like the left-hand panel of Fig. 1. A tier with two factors and no nesting looks like the left-hand panel of Fig. 2 in Section 2.3. We always choose factor names within a tier to have different initial letters, so that a tier with two factors, one nesting the other, can be shown in the abbreviated form demonstrated in the right-hand panel of Fig. 1. Where two or more factors nest another, the nesting factors are shown in a list, as in the right-hand panel of Fig. 4.

It is important to realize that, as discussed by Brien (1983), the nesting relationships reflect those for the design employed, not the inherent relationships between the factors. While in many instances the design employed should respect the inherent relationships, there will be circumstances in which it is desirable that it does not. For example, a randomized complete block design might be used on a rectangular array of plots because the experimenter knows from previous experience that it is almost certain that there will be no trend in one direction. In these circumstances there will be a nesting relationship between the factors indexing the rows and columns of the plots. The continuation of Example 1 will provide another case.

In general, we have in mind that each set of objects is a structured set (Bailey, 1989).

The structure in our examples is provided by the partitions of the set of objects which are defined by the intrinsic generalized factors. When the structure on the set is a poset block structure (Bailey, 1991, 1996), which is a slight generalization of the simple orthogonal block structures introduced by Nelder (1965a), the structure is defined *unambiguously* by some factors, their numbers of levels, and some nesting relationships between them: therefore it can be defined succinctly in a panel. A structure is a poset block structure when the number of levels in the generalized factor comprised of all factors in the tier is the product of the numbers of levels of the (original) factors. All sets for the examples we present are poset block structures. Randomization diagrams can also show the randomization of some treatment structures that are not PBSs; Brien (1992, Section 4.2.3.2) gives an example.

Suppose that Υ is a set of treatments and Ω a set of observational units. The allocation of treatments to units is actually a function f from Ω to Υ: the treatment on observational unit ω is just $f(\omega)$. The usual mathematical notation for this is $f\colon\Omega \to \Upsilon$. Nevertheless, because we speak of randomizing the treatments to the observational units, we use the *opposite* direction for the arrows in our randomization diagrams. We could have chosen either convention, but we must be consistent. With our convention, if a randomization involves two sets of objects of different sizes then the arrow must point to the tier on the larger set, which is the one whose permutations form part of the randomization. We call the tier for the permuted set of objects the *unrandomized* tier, the other the *randomized* tier, for consonance with the common idea that treatments are randomized while observational units are not.

2.3. *Other possible features of a single randomization*

Example 2 (a factorial design): a variant of Fig. 1 shows a randomized complete block design with factorial treatments. Fig. 2 illustrates a poultry-feeding experiment where treatments are randomized to eight cages within each of four rooms. The treatments are all combinations of four quantities of protein with two sources of protein. The two lines meeting at the black circle indicate that all eight levels of the generalized factor Quantities $\wedge$ Sources are randomized to eight Cages within each Room.

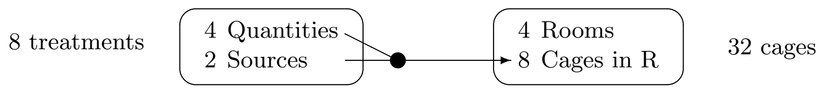

Fig. 2. Randomization diagram for a factorial design in complete blocks

White (1975) discusses the difference between three methods of randomization that are superficially similar. Those with two tiers differ from both the foregoing, as is succinctly shown on randomization diagrams in Fig. 3.

Thus far, every randomization diagram has contained a single arrow from one tier to the other. However, there may be two or more arrows for designs such as a split-plot or criss-cross. Fig. 4 gives the randomization for an extended split-plot design from Mead (1988, page 382), while Fig. 5 gives that for a criss-cross experiment from Clarke and Kempson (1997, page 141). Even though there are two arrows in each example, each involves only a single randomization from one tier to another because they can be achieved in a single permutation. Such randomizations are discussed further in Sections 4.3 and 8.5.

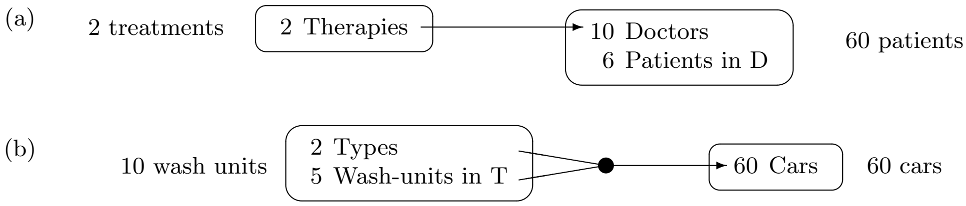

Fig. 3. (a) example (1) and (b) example (3) from White (1975)

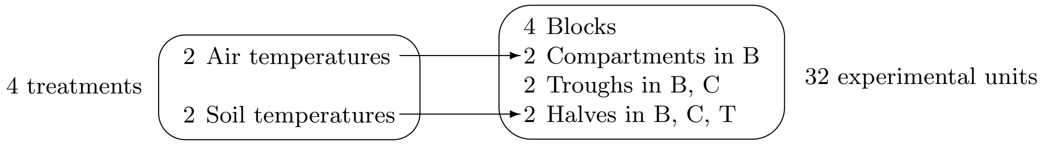

Fig. 4. Randomization diagram for a split-split plot design with two treatment factors

If each factor in the first tier is randomized to a single factor in the second tier then no further information is needed to describe the design: moreover, the design must be orthogonal. Sometimes, however, a factor may be randomized to the combinations of two or more unrandomized factors. In this case, the convention we use is that, if t treatments are randomized to the st combinations of unrandomized factors, then each treatment occurs s times among the combinations. There are essentially four distinct cases. In the first two, where the factor relationships are taken into account, a special systematic design—such as a Latin square, a Youden square, an incomplete-block design or semi-Latin rectangle—is used. Thus more information about the design is needed, and this is indicated by an open circle with a 'perpendicular' sign $\perp$ in it, if the resultant allocation is orthogonal, such as a Latin square, or an open circle otherwise, e.g. for an incomplete-block design. The former is shown on the right-hand side of Fig. 7 in Section 4.1 and Fig 14 in Section 4.3, the latter in Fig. 25 in Section 7. In the third case, where the factor relationships are ignored, the unrandomized factor is completely randomized to the levels of a generalized factor formed from the unrandomized factors. This is indicated in the diagram by a black circle preceded by an arrow and with lines leading from it to all the relevant unrandomized factors. Fig. 20 in Section 6 gives an example of this where the unrandomized factors are nested. In the fourth case, pseudofactors are needed to accomplish the randomization: this is described in Example 5 in Section 4.2.

One further possible feature of a single randomization, namely consonance, is deferred until Section 4.1.

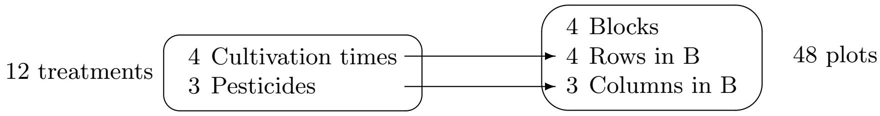

Fig. 5. Randomization diagram for criss-cross design

3. Three tiers

Example 1 (continued) (a two-phase sensory experiment): in the second, or evaluation, phase of this experiment, wines made from the produce of the field trial are evaluated. The produce from each plot is separately made into wine which is evaluated at a tasting in which several judges are given the wines over a number of sittings. One wine is presented for scoring to each judge at a sitting and each wine is presented only once to each judge. The order of presentation of the wines is randomized for each judge.

The example involves two randomizations: treatments to plots and plots to evaluations. The sets of objects are evaluations, plots and treatments and there are three tiers. Fig. 6, which conveniently summarizes both randomizations in a single diagram, also gives the sets of objects, their tiers and the factor nesting. Although Judges and Sittings are inherently crossed on the set of evaluations, the experimenter chose to ignore this inherent relationship, perhaps because he thought that Sittings was unlikely to be an important source of variation. Sittings is nested within Judges and randomization is according to the group $S_{bt} \wr S_j$.

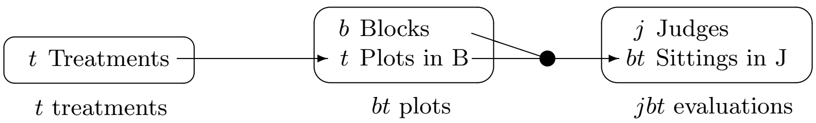

Fig. 6. Randomization diagram for example 1

The basic feature of three-tiered experiments is that they involve multiple randomizations. We identify six types: composed, coincident, independent, double, r(andomized)-inclusive, and u(nrandomized)-inclusive multiple randomizations. Each type consists of two randomizations in each of which one tier is randomized to another tier, there being three tiers. The first four types we group together as randomizations that are independent of order: they are discussed and exemplified in Section 4. For these types, the order in which the two randomizations are performed is arbitrary: indeed, each can be performed without any knowledge of the uninvolved, third tier. The two inclusive multiple randomizations are described and illustrated in Section 5.

4. Two Randomizations Independent of Order

4.1. Composed randomizations

Composed randomizations occur when there is an initial randomization of a set of objects Γ to another set of objects Υ, and then Υ is randomized onto a third set of objects Ω. Thus a permutation of Υ is chosen to randomize Γ and a permutation of Ω is chosen to randomize Υ. The permutations may be chosen in either order: the outcome of one does not affect choice of the other. The randomizations are composed in the sense that the randomization of Γ onto Ω can be obtained by first randomizing Υ onto Ω and then randomizing Γ onto the previously randomized Υ. In other words, the function which maps each element of Ω to the element of Γ which is randomized to it is the composite of the function mapping each element of Ω to the element of Υ randomized to it with the function mapping each element of Υ to the element of Γ randomized to it.

In Example 1 there are two randomizations—treatments to plots and plots to evaluations—and they are composed. No account is taken of the randomized factor from the first phase when doing the randomization of the second phase. Nonetheless, the two randomizations taken together have the effect of randomizing treatments to evaluations. Example 15 in Section 7 is a more complicated two-phase sensory experiment involving composed randomizations.

The literature contains several examples of experiments with composed randomizations: two-phase experiments are given by McIntyre (1955), Cox (1958, Example 3.2, as discussed in our Example 4), Brien, May and Mayo (1987), Wood, Williams and Speed (1988); single-stage experiments are described by White (1975, example (2)), Preece (1991) and Brien and Demétrio (1998, as discussed in our Example 3). At the Joint Statistical Meetings in New York in 2003, T. B. Bailey described the two-phase experiment shown in Fig. 7: the design in the second phase consists of a pair of 6×6 Latin squares.

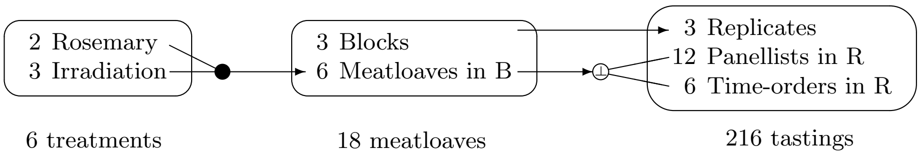

Fig. 7. Composed randomizations in a two-phase experiment

Example 3 (a continuous grazing experiment): Brien and Demétrio (1998) discuss the analysis of continuous grazing trials based on randomized complete block designs. The sets of objects for such experiments, when an animal variable is to be analysed, are animals, plots and treatments. The animals are divided into classes according to breed or weight. These experiments involve two randomizations—treatments to plots and plots to animals—and they are composed, as shown in Fig. 8, which is equivalent to Fig. 6. In spite of this, the present example is not a two-phase experiment: it is single stage.

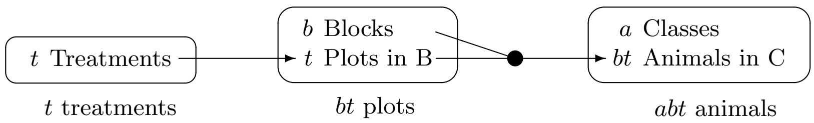

Fig. 8. Composed randomizations in a single-stage experiment in Example 3

Example 4 (cotton fibres): Cox (1958, Example 3.2) describes a two-phase experiment that illustrates a number of issues. It consists of a field phase in which cotton is produced in a field trial and a testing phase in which the cotton is tested for strength. In the field phase five quantities of potash are applied to a cotton crop, in three blocks of five plots. Thus there is one set of five treatments (the levels of Potash) and one set of fifteen plots (the levels combinations of Blocks with Plots). In the testing phase "a number" of fibres of cotton are sampled from each plot and tested for strength. For definiteness suppose that two fibres per plot are tested, so that the set of fifteen plots is replaced by a set of 30 fibres. Thus there are two randomizations: treatments are randomized to fibres and the fibres are randomized to tests. The randomization of treatments to fibres is settled, but there are a number of possibilities for randomizing the fibres to tests. The randomization diagrams will help the experimenter to choose between different ways of performing this second randomization.

Cox recommends that all fibres from each block should be tested by a single operative. If there are three operatives then there is a third tier on 30 tests, shown in the right-hand panel of Fig. 9. The pair of arrows on the right show Cox's recommended randomization: Blocks are randomized to Operatives, while the ten fibres tested by each operative are in independent random orders.

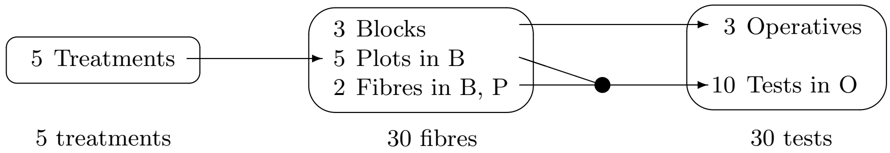

Fig. 9. Composed randomizations in the Example 4: Plan A

Alternatively, if there are two operatives, each could test 15 fibres, one from each plot. Fig. 10 is the randomization diagram for both phases. There is a small difficulty in drawing this. Fibres are nested in Blocks ∧ Plots, but they are randomized to Operatives, which nests Tests, to which Blocks∧Plots is randomized. We say that the nesting is not *consonant* with the randomization. Thus we must *either* have arrows crossing *or* write a nested factor above a non-nested one in one tier. Since the nesting relationships are shown clearly in words within each tier, we prefer the latter.

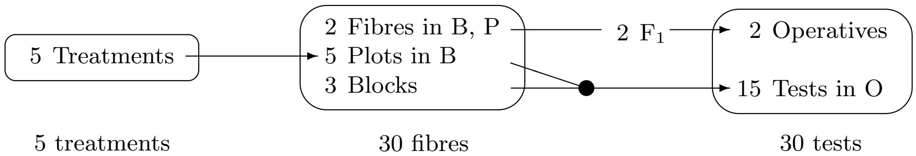

Fig. 10. Composed randomizations in Example 4: Plan B

The lack of consonance forces us to introduce a two-level pseudofactor F_1 (Section 8.2) that indexes the two groups of 15 fibres; it is not nested in anything. We use the convention that pseudofactors have the same letter as the (real) factor and are distinguished by subscripts. The pseudofactor has no inherent meaning in the middle tier, it representing arbitrary differences between Fibres. Consequently, the pseudofactor is not incorporated into the panel for the tier. To avoid confounding Operatives with any systematic difference between the fibres, it is necessary to either randomly label the levels of Fibres within each plot or to randomly assign the fibres from each plot to the two groups indexed by F_1.

A third possibility, using six operatives each testing one fibre per plot of one block, is shown in Fig. 11.

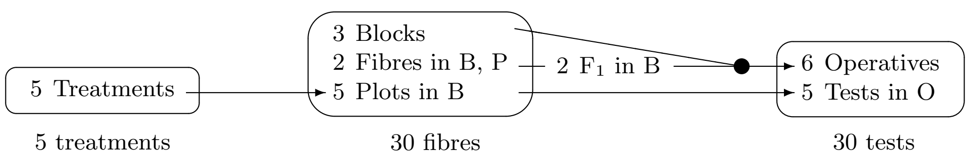

Fig. 11. Composed randomizations in Example 4: Plan C

4.2. Coincident randomizations

Coincident randomizations occur when Γ and Υ are both randomized to Ω, in such a way that Γ and Υ are randomized indirectly to each other. That is, restrictions on the randomizations are provided only by the factors on Ω. Thus two permutations from the same group on Ω are chosen independently: one is used to randomize Γ to Ω and the other to randomize Υ to Ω, and the order in which they are applied is arbitrary.

Coincident randomizations are needed when it is impossible to observe not only all combinations $\Omega \times \Gamma$ and all combinations $\Omega \times \Upsilon$, but also all combinations $\Gamma \times \Upsilon$. It is therefore necessary for some factor(s) on Ω to be involved in both randomizations. Some care needs to be taken to keep some control over the ensuing confounding between Γ and Υ.

Example 5 (a plant experiment): a plant experiment is conducted to investigate five varieties and two spray regimes. Twelve seedlings of each variety are put into individual pots; these are randomly assigned to six benches on each of which there are 10 positions so that there are two seedlings of each variety on each bench. The two spray regimes are randomly assigned to the benches so that each is applied to the pots on three benches. The height gain after six months is measured for each seedling.

The sets of objects for this experiment are positions, seedlings and regimes. The tiers and factor nesting are shown in Fig. 12. This single-stage experiment involves two randomizations: seedlings to positions and regimes to positions.

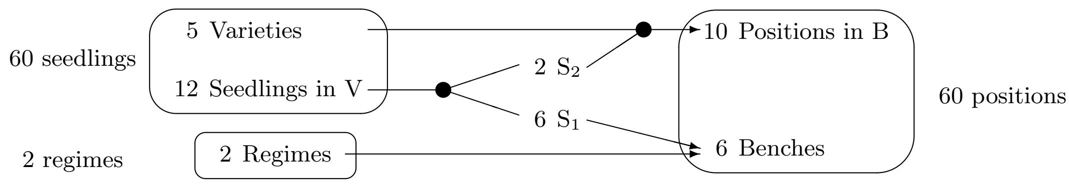

Fig. 12. Coincident randomizations in Example 5

In randomizing seedlings to positions, the factor Seedlings (in Varieties) is randomized to two factors from the positions tier, Benches and Positions (in Benches), and the randomization to Benches is not consonant. For this, the seedlings within each variety are split into six sets of two. It is convenient to represent the formation of sets of seedlings by using pseudofactors S_1 and S_2 for Seedlings, with six and two levels respectively, so that Seedlings $= S_1 \wedge S_2$. The lines and black circle between Seedlings and the pseudofactors in the diagram portray this splitting of Seedlings into the two pseudofactors. The use of pseudofactors here is consistent with the concept of generalized factors. In Example 2, Quantities $\wedge$ Sources = Treatments, while here $S_1 \wedge S_2$ = Seedlings. In both cases the real factors are those listed within the panels. Now Varieties $\wedge$ S_2 is randomized to Positions in Benches. On the other hand, S_1, taken across all levels of Varieties, is randomized to Benches. As in Example 4 (Plan B), it is necessary to randomly permute the 12 levels of Seedlings within each Variety, so that neither S_1 nor S_2 corresponds to any inherent feature such as vigour.

The randomizations are summarized in Fig. 12. They are coincident because not every combination of Varieties $\wedge$ Seedlings and Regimes can occur: the confounding between Regimes and part of S_1 is indicated by the fact that both are randomized to Benches.

Coincident randomization involves two randomizations to the same unrandomized factors; this is indicated in the diagram by either two arrowheads or sets of lines from two circles going to the same factor(s). Miller (1997) has a two-stage experiment in which different treatment factors are randomized at each stage, so that two of the randomizations are coincident. Mee and Bates (1998) extend this to three or more stages. However, two lines or arrows to a factor do not always indicate coincident randomizations: see Example 7 (Section 4.3).

4.3. Independent randomizations

Independent randomizations happen when both Γ and Υ are randomized to Ω, in such a way that $\Gamma \times \Upsilon$ is observed and there is no random association between Γ and Υ. Generally, these randomizations, each using an allowable permutation of Ω, could be performed in either order, although in some cases there will be practical reasons for performing one before the other.

Independence means that the two separate systematic designs may be combined into a systematic design for $\Gamma \times \Upsilon$. Moreover, the randomizations can often be achieved in a single randomization of $\Gamma \times \Upsilon$ to Ω. They cannot be achieved in this way if the randomizations are at different times. However, Schoen's (1997) experiment demonstrates that a multi-stage splitting experiment can involve independent randomizations that are achievable in a single randomization. Example 12 in Section 6 shows another way in which independent randomizations cannot be achieved in a single randomization.

Strictly speaking, the set Ω is the same for a pair of independent randomizations. Thus, for the split-plot experiment, the subplots are available for both randomizations, although nothing is randomized to different subplots in a main plot in the first randomization. It is, however, sometimes convenient to regard Ω as changing between randomizations. Using this second method for a split-plot experiment, we would first randomize to main plots, there being no subplots identified yet. Subsequently, subplots would be added, changing the objects from main plots to subplots, and we then randomize to subplots. Clearly, these two randomizations cannot be combined into one and, in practice, that for main plots would be done before that for subplots. In general, independent randomizations done in this way firstly randomize Γ to a set Δ, which is uniquely indexed by its tier $\mathcal{F}_\Delta$. Then Δ is expanded to the set Ω by replacing each object in Δ by a constant number of objects in Ω. The tier of factors on Ω must include $\mathcal{F}_\Delta$ but none of the extra factors should nest any in $\mathcal{F}_\Delta$. Then Υ is randomized to Ω, using a systematic design in which all the objects in Υ occur equally often on each level of the generalized factor formed from all the factors in $\mathcal{F}_\Delta$.

Example 6 (a superimposed experiment using split plots): a randomized complete block experiment with b blocks is set up to investigate the yield differences between r rootstocks for orange trees, each plot containing t trees. After several years of running this initial experiment, it is decided to incorporate t fertilizer treatments by randomizing them to the t trees in each plot.

The sets of objects for this experiment are trees, rootstocks and fertilizers. There are two randomizations: rootstocks to plots in the initial experiment and fertilizers to trees in the revised experiment. The performance of these randomizations being separated in time, they cannot be achieved in a single randomization. However, the two randomizations are independent—they involve two disjoint subsets of factors in the trees tier: see

Fig. 13. Therefore the factors Rootstocks and Fertilizers are not associated by randomization. However, these last two tiers are not combined because, as a result of their separate randomizations, the factor in one tier has a different status in the randomization from that in the other tier. Further, the randomization to plots precedes the randomization to trees, although theoretically they could be done in either order.

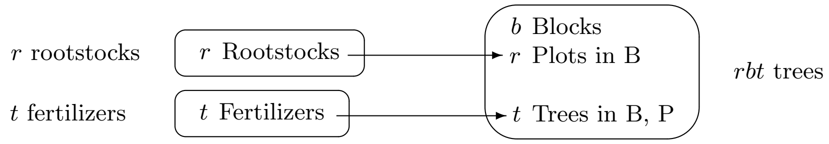

Fig. 13. Independent randomizations in Example 6

Example 7 (an experiment in human-computer interaction): an experiment is run to compare two methods in which a person can draw a map in a computer file: using a mouse and using a stylus. Twelve subjects are recruited to the experiment, and are randomized to six days and to tests within days. This can be accomplished by using pseudofactors for Subjects to randomly divide the subjects into groups of two. Each day one subject uses Room A in the morning (Period 1) followed by Room B in the afternoon (Period 2), while the other uses the rooms in the other order. The two Methods are randomized to Days $\wedge$ Rooms using three 2×2 Latin squares. The sets of objects for this experiment are tests, subjects and methods. There are two independent randomizations that are reducible to a single randomization by combining two of the sets into a set of 24 subject-by-method combinations. Fig. 14 summarizes the randomization.

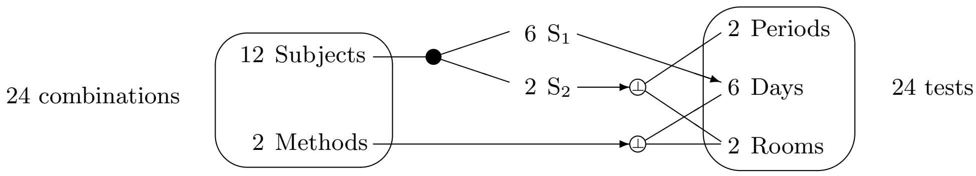

Fig. 14. Independent randomizations in Example 7

Although some lines go to the same factor, this experiment does not involve coincident randomizations. This is because the systematic designs involved, indicated by the open circles with $\perp$ signs, randomize the (pseudo)factors to different terms. For example, while pseudofactor S_1 is randomized to Days, Methods is randomized to Days $\wedge$ Rooms such that Methods is not confounded with Days.

4.4. Double randomizations

Double randomizations occur when Γ is randomized to both Υ and Ω. As a result, elements of Υ and Ω are connected by being assigned the same element of Γ. A permutation is chosen for each of Υ and Ω and each applied independently to Γ. It appears that this is a degenerate form of randomized-inclusive randomization, to be discussed in Section 5.1, in that it applies only when one of the randomizations involves two sets of objects of the same size. The order of the two randomizations is also arbitrary in these circumstances. In Example 8 there is only a single replicate of each treatment, so that the sizes of the two sets of objects in one of the randomizations are equal.

Example 8 (an improperly replicated rotational grazing experiment): as 't Mannetje, Jones and Stobbs (1976) note, one way to reduce herd size in grazing trials is to use a single herd to graze the replicates of each treatment in turn. For example, an experiment is conducted to investigate the effects of three levels of pasture availability on the weight gain of cattle. The 12 combinations of three levels of availability and four rotations are applied completely at random to 12 paddocks. Also, the levels of availability are assigned completely at random to 15 animals so that each level of availability is assigned to five animals. The five animals are then grazed together in sequence on the four paddocks assigned to that level of availability; the sequence of four paddocks is determined by the order in which the rotations are assigned to them.

The sets of objects for this experiment are paddocks, observational units (cattle over time) and treatments. The randomization diagram is in Fig. 15. The factor Rotations occurs in two tiers. It must be included in one so that the observational units are uniquely indexed by the set of factors in this tier. It occurs in the other tier because Availability$\wedge$Rotations is randomized to Paddocks. It is not possible to have the Rotations in the tier of observational units randomized to Paddocks as there is no viable systematic design. This experiment involves double randomizations because Availability is randomized twice: these are shown by two arrows which start from the same factor but go to different tiers. There is no arrow between the two Rotations in the different tiers because they are identical.

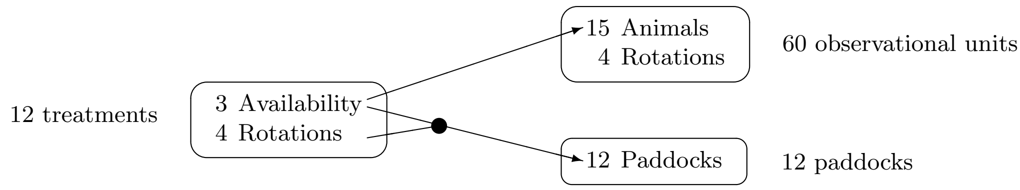

Fig. 15. Double randomizations in Example 8

5. Two Inclusive Randomizations

The multiple randomizations described so far can be performed in either order because each single randomization uses information from only two tiers. For the remaining two types of multiple randomization, neither of these features applies: information from the first randomization must be used in carrying out the second; consequently, it must be done first. We call these types *inclusive* because factors from both tiers of the first randomization are explicitly involved in determining the design for the second randomization.

In one case, described in Section 5.1, the first randomization randomizes Γ to Υ and the second randomizes Υ to Ω using a systematic design which incorporates information about the factors on Γ. We call this multiple randomization *randomized-inclusive* because all the factors from the first randomization ($\mathcal{F}_\Gamma \cup \mathcal{F}_\Upsilon$) form a pseudotier that is randomized in the second randomization. The 'r' indicates that the factors in the pseudotier are the randomized factors in the second randomization.

In the other case, described in Section 5.2, the first randomization randomizes Υ to Ω and the second randomizes Γ to Ω, using a systematic design which incorporates information about the factors on Υ. We call this multiple randomization *unrandomized-inclusive* because all the factors from the first randomization ($\mathcal{F}_\Upsilon \cup \mathcal{F}_\Omega$) form a pseudotier that becomes the set of unrandomized factors in the second randomization. The 'u' indicates that the factors in the pseudotier are the unrandomized factors in the second randomization.

Both of these types of inclusive randomization differ from the multiple randomizations of Section 4. However, an randomized inclusive multiple randomization has in common with double and composed randomizations that one of the three tiers is randomized twice. Given that firstly Γ is randomized to Υ: if the second randomization to Ω uses information from Γ only, then the multiple randomization is double; if it uses information from Υ only, then the multiple randomization is composed; if it needs information from both Γ and Υ, then the multiple randomization is randomized inclusive. A unrandomized inclusive multiple randomization has in common with coincident, independent and composed randomizations that there is a tier which is unrandomized twice. Given that Υ is first randomized to Ω and that Γ is also randomized: if the unrandomized set of objects for the second randomization is just Ω, then the multiple randomization is coincident or independent; if Γ is randomized directly to Υ and hence indirectly to Ω then the multiple randomization is composed; if the second randomization uses information from both Υ and Ω then the multiple randomization is unrandomized inclusive.

5.1. Randomized-inclusive randomizations

The initial randomization of a pair of randomized-inclusive randomizations randomizes Γ, with tier $\mathcal{F}_\Gamma$, onto Υ, with tier $\mathcal{F}_\Upsilon$. The second randomization uses a new pseudotier whose set of objects is Υ and whose set of factors is $\mathcal{F}_\Gamma \cup \mathcal{F}_\Upsilon$. This is randomized to Ω, with its tier $\mathcal{F}_\Omega$. Thus a permutation of Υ is chosen to randomize Γ; a permutation of Ω is chosen to randomize Υ. The choice of the systematic plan allocating Υ to Ω (pre-randomization) is constrained by the outcome of the first randomization.

Randomized-inclusive randomizations are needed when there is a factor F from the middle tier $\mathcal{F}_\Upsilon$ such that

(a) the levels of at least one (generalized) factor from $\mathcal{F}_\Gamma$ are randomized to F in the first randomization;
(b) F is randomized to the levels of more than one factor from $\mathcal{F}_\Omega$ in the second randomization;
(c) the systematic design for the second randomization is not balanced for factor F (in the sense of "balanced block design").

Denote by $\mathcal{G}_\Gamma$ the set of factors randomized to F. Because of (b) and (c), a set $\mathcal{H}_\Upsilon$ of pseudofactors is needed for F. Because of (a), a similar set $\mathcal{H}_\Gamma$ may be needed for factors from $\mathcal{G}_\Gamma$. Of course, the pseudofactors in $\mathcal{H}_\Upsilon$ need to parallel either the factors from $\mathcal{G}_\Gamma$, or the pseudofactors $\mathcal{H}_\Gamma$.

Example 9 (a two-phase wheat variety trial) (this example is based on a problem described to us by K.A. Haskard, BiometricsSA): in the field phase of this experiment 49 lines of wheat are investigated using a randomized complete block design with four blocks. The produce of each plot is analysed using a gas chromatograph in which seven samples can be processed per run. A 7×7 balanced lattice-square design with four replicates is used to assign the blocks, plots and lines to four intervals in each of which there are seven runs at which samples are processed at seven consecutive times.

The sets of objects for this experiment are analyses, plots and lines. Fig. 16 shows this experiment's two randomizations: lines to plots and plots to analyses.

Three aspects of this experiment prevent the use of composed randomizations:

(a) lines are randomized to plots within blocks in the first randomization;
(b) plots within blocks are randomized in the second randomization to the levels of, not a single analyses factor, but two different analyses factors, and their combinations;
(c) there are two efficiency factors, 0 and 1, for plots within blocks with each analyses term to which it is randomized.

These obstacles could be overcome, and composed randomizations employed, by assigning the 49 plots within a block completely at random to the 49 run-time combinations within an interval. However, then it would be inefficient to isolate the Runs-within-Intervals and Times-within-Intervals effects in the analysis of variance, Lines being hopelessly confounded with these two terms. So we need to choose different pseudofactors for Plots within Blocks to randomize to different analyses factors. One pseudofactor P_1 with 7 levels is randomized to Runs within Intervals and another, P_2, to Times within Intervals. We ensure that $P_1 \wedge P_2 =$ Plots so that all 49 Plots within a block are randomized to the 49 Runs by Times combinations within an interval. Pseudofactors within each block are selected so that Lines are balanced with the three terms involving Runs and Times. Thus it is necessary to know which Lines are associated with which Plots within Blocks—i.e. to know the outcome of the first randomization. Here a balanced lattice-square design is used to specify that the plots with certain lines should be processed at the same run in a particular interval or at the same time in a particular interval. This grouping of the lines for randomizing in a particular interval automatically groups the plots in the block assigned to that interval. A pair of Lines pseudofactors is set up for each interval, one to indicate those lines that occur in the same run in that interval and the other to indicate those that occur at the same time in that interval. Label the pairs: $(\mathrm{L}_1, \mathrm{L}_2)$, $(\mathrm{L}_3, \mathrm{L}_4)$, $(\mathrm{L}_5, \mathrm{L}_6)$ and $(\mathrm{L}_7, \mathrm{L}_8)$.

The randomization diagram for this experiment has two new features. The dashed oval shows the pseudotier created by the first randomization. Diamonds show which factor(s) or pseudofactor(s) from the first tier have been used to define the pseudofactors in the second tier, pseudofactors which cannot be identified until the first randomization has been done. Here each diamond, with lines going towards it, indicates that four of the Lines pseudofactors are used to define the corresponding Plots pseudofactor, which is then randomized to Runs or Times within Intervals, as appropriate.

The permutation group in the first randomization is $S_{49} \wr S_4$ on plots. In the second randomization it is $(S_7 \times S_7) \wr S_4$ on analyses. Thus the effect of the first randomization is to constrain, not the second permutation group, but the choice of initial systematic plan allocating the plots to analyses.

Like Examples 1 and 12, this is a two-phase experiment, involving different units in its field and laboratory phases. However, it differs from those examples in that it involves randomized inclusive, rather than composed, multiple randomizations because factors from both tiers of the first randomization are explicitly included in the randomized factors for the second randomization.

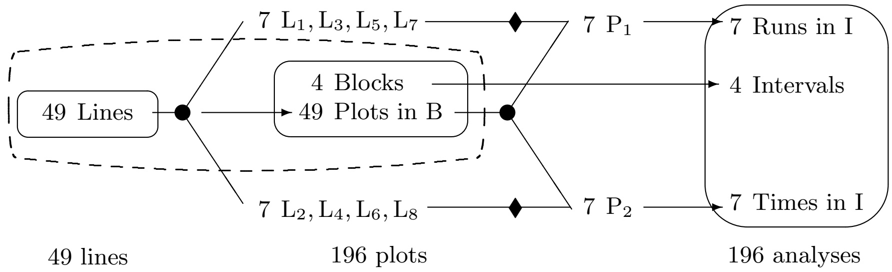

Fig. 16. Randomized-inclusive randomizations in Example 9

Example 8 (continued) (an improperly replicated rotational grazing experiment): we have previously presented this example as a case of double randomization. Here we show that it can be regarded as randomized-inclusive randomization.

The sets of objects for the experiments are as before. Here we identify the two randomizations as treatments to paddocks and paddocks to observational units (cattle over time). The randomizations are randomized inclusive because the second one randomizes the paddocks with the same level of Availability to the same Animal—this requires knowledge of the assignment of Availability to Paddocks. The randomization diagram in Fig. 17 includes pseudofactors $\mathrm{P_A}$ and $\mathrm{P_R}$ for Paddocks that are determined by the levels of Availability and Rotations, respectively, assigned to Paddocks.

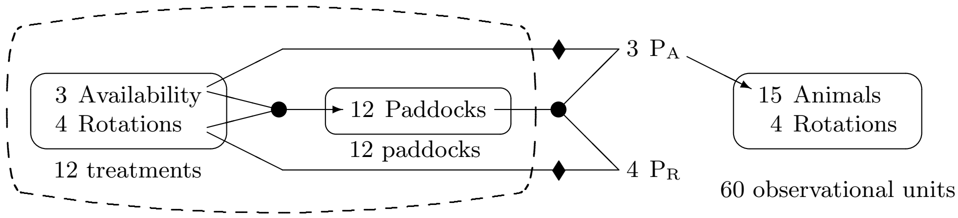

Fig. 17. Randomized-inclusive randomizations in Example 8

The simpler representation of the randomization for this example, given in Fig. 15, is possible because of the equal numbers of paddocks and treatments. Basically, the pseudofactors from Fig. 17 are omitted. The simpler diagram has Availability directly randomized to both Paddocks and Animals and it would seem that all that is involved is a double randomization of Availability, along with the randomization of Rotations to Paddocks. However, the randomization of Availability to Animals will also result in the randomization of Paddocks to Animals and this is left implicit in Fig. 15.

5.2. Unrandomized-inclusive randomizations

For unrandomized-inclusive randomizations, the unrandomized factors for the second randomization are all those from the first randomization. So a unrandomized-inclusive randomization takes place when there is an initial randomization of Υ, with tier $\mathcal{F}_\Upsilon$, onto Ω, with tier $\mathcal{F}_\Omega$, after which Γ is randomized to the new pseudotier $\mathcal{F}_\Upsilon \cup \mathcal{F}_\Omega$ on the set Ω. Thus a permutation of Ω is chosen to randomize Υ; a second permutation of Ω, subject to restrictions involving $\mathcal{F}_\Upsilon \cup \mathcal{F}_\Omega$, is then chosen to randomize Γ.

Table 1. Virus Treatment for each Block-Rootstock combination

		Rootstocks									
		1	2	3	4	5	6	7	8	9	10
	I	A	B	A	C	D	C	B	E	E	D
Blocks	II	D	E	B	D	E	A	C	C	A	B
	III	E	A	C	E	B	D	D	B	C	A

Example 10 (a superimposed experiment using a row-column design): Freeman (1959) describes a cherry experiment in which a large number of rootstocks had been tested using a randomized complete block design for 20 years. The trees from 10 rootstocks in three blocks were then to be used for investigating five virus treatments. The five virus treatments were assigned using the extended Youden square in Table 1.

The sets of objects for this experiment are trees, rootstocks and treatments, with tiers shown in Fig. 18. There are two randomizations: rootstocks to trees in the initial experiment and virus treatments to trees, taking into account the rootstocks, in the revised experiment. The two randomizations are unrandomized inclusive in that the unrandomized factors for the second one are those from both the rootstocks and trees tiers of the first one. The dashed oval once again shows the new pseudotier created by the first randomization.

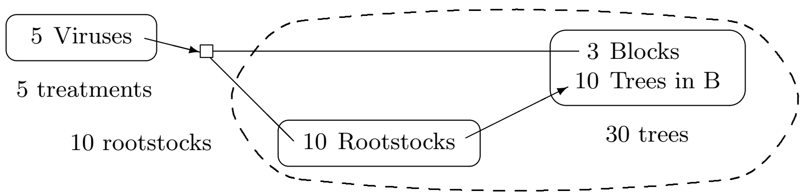

Fig. 18. Unrandomized-inclusive randomizations in Example 10

The square after the arrow indicates that viruses are randomized to those combinations of Blocks and Rootstocks that occur as a result of the first randomization. The square shows combinations of levels of factors from different tiers, in contrast to circles that show combinations from the same tier. Like an open circle, the open square indicates that a special, nonorthogonal design was used. The group for the first randomization is $S_{10} \wr S_3$, while that for the second is $S_{10} \times S_3$, which is a subgroup of the former. In contrast to Example 9, both the group and the systematic plan for the second randomization are constrained by the result of the first. This double constraint is reflected in the number of plans which the designer must write out: both the result of the first randomization and the randomized version of the constrained systematic plan in Table 1 are needed to construct the plan for the orchard worker to follow.

5.2.1. Incoherent unrandomized-inclusive randomizations

Example 10 demonstrates an important principle of unrandomized-inclusive randomizations: factors which are nested for one randomization can apparently become crossed for a subsequent randomization. This is permissible because crossing is more restrictive than nesting. However, when a randomization to crossed factors has been appropriately restricted, all subsequent randomizations should respect this crossing. Having decided that

Table 2. Design from Cochran and Cox (1957): s_0–s_3 are soil treatments; numbers denote leaf treatments

	Bench I				Bench II				Bench III			
	s_3	s_2	s_1	s_0	s_0	s_2	s_1	s_3	s_3	s_0	s_2	s_1
Top layer	2	1	0	1	0	0	1	1	0	2	2	2
Middle layer	0	0	1	0	2	2	2	2	1	1	1	0
Bottom layer	1	2	2	2	1	1	0	0	2	0	0	1

a factor is likely to be important and so ensured that its effects can be separated out, it is inconsistent to then, in a subsequent randomization, loosen the restrictions and treat this factor as nested. If any proposed pair of unrandomized-inclusive randomizations breaks this principle then we deem them *incoherent*.

This principle does not apply to all replacements of a crossed with a nested relationship. It does not apply when an inherent crossed relationship is replaced with a nested one in the design, as discussed in Section 2.2. In such cases, the randomization is perfectly coherent.

In Example 10 the relationship on Ω is initially Trees nested within Blocks, so that the only non-trivial partition on Ω is into the three blocks: this is preserved by the group $S_{10} \wr S_3$. After the first randomization, each tree can be uniquely identified by the levels of Blocks and Rootstocks, so Rootstocks can effectively replace Trees as a factor. However, Rootstocks are crossed with Blocks: there are now two non-trivial partitions on Ω, into the three blocks and into the ten rootstocks: the smaller group $S_{10} \times S_3$ preserves them both.

Suppose that the designer of the superimposed experiment decides to ignore the inherent crossing of Blocks and Rootstocks and randomizes Viruses to Blocks in Rootstocks in a balanced incomplete block design. Then the permutation group for the second randomization is $S_3 \wr S_{10}$, and the crossed relationship from the first randomization is changed a into a nested relationship in the second. The randomizations are incoherent. The experimenter thinks that it is important to allow for the Blocks in the first randomization, but over-rides this in the second. In more complicated situations it is not always so easy to see that a proposed unrandomized-inclusive randomization is incoherent.

Example 11 (split plots in a row-column design): Cochran and Cox (1957, Section 7.33) present an experiment in which the subplots are arranged in a row-column design. For clarity, we give definite names to the factors, using one of their suggested applications. Four soil treatments are randomized to four pot plants on each of three benches. Three leaf treatments are applied to each plant, one each to the upper, middle and lower leaves. These are randomized independently within each soil treatment, using a Latin square in which the rows are layers and the columns are benches, giving the plan in Table 2. The plants are the main plots and the leaves the subplots. The experiment is unusual in that the subplot treatments are randomized within the levels of the main-plot treatments.

The sets of objects for this experiment are leaves, soil treatments and leaf treatments. There are two randomizations: soil treatments to plants, shown at the top of Fig. 19, and leaf treatments to leaves, taking account of soil treatments, shown at the bottom of Fig. 19. They are unrandomized inclusive because the unrandomized factors for the second randomization must include factors from the both tiers of the first.

It is clear that Plants is nested in Benches and that Layers is inherently crossed with both. The experimenter is explicitly concerned about “a regular gradient of susceptibility

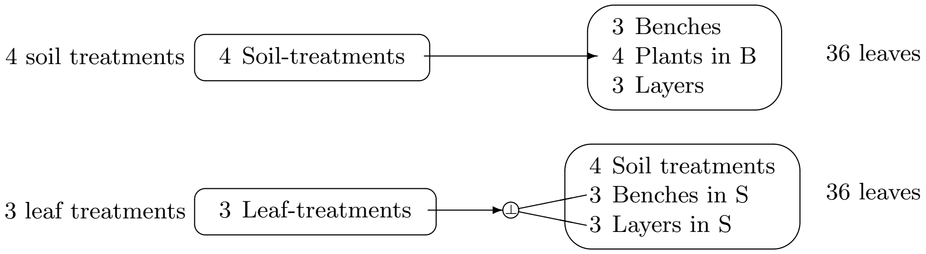

Fig. 19. Incoherent randomizations in Example 11

down the plant" which is constant across benches. Therefore the experimenter takes notice of this nesting and crossing, and the group of permutations for the first randomization is $(S_4 \wr S_3) \times S_3$.

In the second randomization, Soil-treatments replaces Plants as an indexing factor, and the group of permutations is $(S_3 \times S_3) \wr S_4$, which is not a subgroup of the first group. Benches nests Plants in the first randomization, but then, in the second randomization, Benches is nested in Soil-treatments, the factor which has been randomized to Plants. This reversal of the nesting between the two randomizations makes them incoherent.

The incoherence is demonstrated in the impossibility of drawing the leaves tier in a consistent manner for the two randomizations. It also leads to difficulties in formulating a mixed model for the data (Section 7): Cochran and Cox (1957) advocate including the intertier interaction Layers $\wedge$ Soil-treatments without commenting on why they omit the more plausible intratier interaction Layers $\wedge$ Benches.

The experiments described by Federer (1975, Example 5.1) and Box and Jones (1992, Table 10) are similarly incoherent.

6. Three or more randomizations

Here we give examples involving more than two randomizations and hence more than three tiers. They also generally involve more than one type of multiple randomization.

With three randomizations there is the possibility of an entirely new form of randomization where the third randomization is at the same time both randomized and urandomized inclusive. Example 13 can be viewed in this light if the first randomization is of treatments to samples, the second of times to pallets and the third of samples to pallets. It does not seem worthwhile to name the other possibilities that can occur, as there are no new principles involved.

Example 12 (a two-phase corn seed germination experiment) (T.B. Bailey, Iowa State University, kindly provided a more complicated version of this experiment): a study to investigate corn seed germination involved a field and a laboratory phase. In the field phase, an experiment to produce corn seed was conducted at three sites; at each site a randomized complete block design of two blocks was run to investigate the differences between three mechanical harvesters. Four samples of seed were harvested from each plot, combined and then divided into 36 lots. In the laboratory phase of the experiment, there were nine containers each with four plates that were used for germinating seed in each of 18 intervals. In each interval, the 36 lots of the seed from one plot were placed on the four plates in the nine containers, the interval in which the seed from a plot was germinated being assigned at random. In each interval, nine temperature-moisture conditions, referred to as nine treatments, were randomly assigned to the nine containers. Thus the inherent crossing of containers with intervals was ignored. The percentage germination of the seeds was recorded for each plate.

This two-phase experiment differs from that in Example 1 in that treatments are introduced in both phases. The sets of objects are plates, lots, harvesters and treatments. The randomization diagram is in Fig. 20. There are three randomizations: harvesters to plots, lots to plates and treatments to plates. The first two are composed and the third is partly coincident with and partly independent of the second. The 36 Lots are randomized to 36 levels of Containers $\wedge$ Plates. This randomization and that of Plots $\wedge$ Sites $\wedge$ Blocks are independent, but can be reduced to single randomization. The randomization of Lots is coincident with that of treatments. On the other hand the randomizations of Plots $\wedge$ Sites $\wedge$ Blocks and treatments are independent, but cannot be reduced to a single randomization.

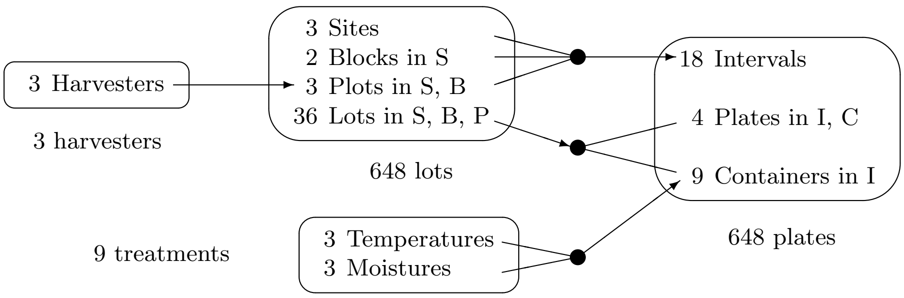

Fig. 20. Composed and coincident randomizations in Example 12

The permutation group used for the first randomization is $S_3 \wr S_2 \wr S_3$ (since Lots are only implicit at this stage) while that for the coincident second and third randomizations is $S_4 \wr S_9 \wr S_{18}$.

The factor Lots could be omitted from the above description, but it is included to emphasize that the 36 lots from a given plot must not be allocated systematically to the 36 plates. A similar remark applies to the factor Seedlings in Example 5.

Example 13 (a two-phase potato storage experiment) (R.W. Payne, VSN International, kindly provided this experiment): this two-phase experiment has field and storage phases. In the field phase four cultivars of potatoes and three fungicides are investigated using a randomized complete block design with three blocks. At the storage phase, the produce from each of the 36 plots is divided into four samples for storage on four pallets, the produce on each pallet being stored for one of four different lengths of time. Altogether, there are 12 pallets on each of 12 benches. The task is to randomize the 144 samples so that the three blocks by four cultivars are randomized to the 12 benches and the three fungicides with their four samples for that bench are randomized to the 12 pallets within each bench. Formally, we identify sets of plots within each block that have the same cultivar and set up a pseudofactor that indexes these sets. Similarly we create a pseudofactor that indexes the sets of plots in each block that received the different fungicides. These two pseudofactors have the same levels as Cultivars and Fungicides, respectively, but differ from them in being nested within Blocks. The desired randomization of the 144 samples to the 144 pallets is achieved by randomizing Blocks $\wedge$ P_C to Benches and Samples $\wedge$ P_F to Pallets in Benches. Finally the four times of storage are randomized to the pallets within each bench-fungicide combination.

Like the two-phase experiment in Example 9, this experiment involves randomized-inclusive randomizations. They are needed because the factor Plots in Blocks:

(a) has treatments randomized to it in the first randomization;
(b) is randomized to Benches and to Pallets in Benches in the second randomization;
(c) has two efficiency factors, 0 and 1, in its randomization to Benches, and also to Pallets in Benches.

This experiment also has in common with the two-phase experiment in Example 12 that treatments are introduced in both phases. The sets of objects are pallets, samples, treatments and times. Fig. 21 shows the three randomizations: treatments to samples, samples to pallets and times to pallets. The first and second are randomized inclusive, and the second and third are unrandomized inclusive. The randomization of samples to pallets involves both treatments and samples factors and so the randomized factors for the second randomization involve both tiers from the first randomization. The randomization of the times is to the pallets within each bench-fungicide combination, so the unrandomized factors for the third randomization come from the treatments and pallets tiers.

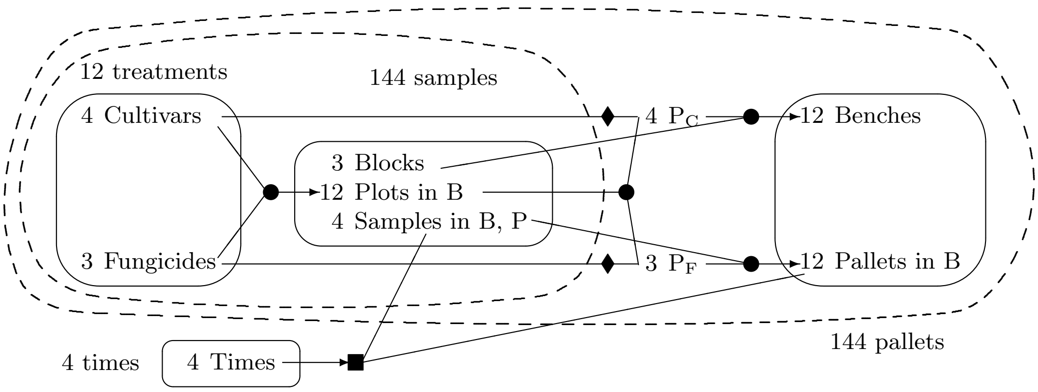

Fig. 21. Unrandomized- and randomized-inclusive randomizations in Example 13

Example 14 (knitted socks): Cox (1958, Example 5.9) discusses four methods of randomizing a two-stage experiment to compare the effects of four chlorination treatments on knitted socks. Each treatment is applied to twelve socks. In the second stage, the socks are tumbled in a machine that simulates normal wear and washing. The machine can take up to twelve socks in each wash. Finally the shrinkage in each sock is measured.

In method I, the socks are randomly divided into four batches of twelve. Treatments are randomized to batches. The third tier consists of 48 rides in the machine split into four washes, one for each treatment. This is shown in Fig. 22(a). It contains a tier for batches with the factor Batches representing possible meaningful differences. It is necessary because each treatment is applied to a whole batch at once, a novel feature of this example.

The diagram also shows the randomization of socks to rides using pseudofactors for Socks. The problem with this method is that Treatments are confounded with Batches that are confounded with Washes and so all three are inseparable.

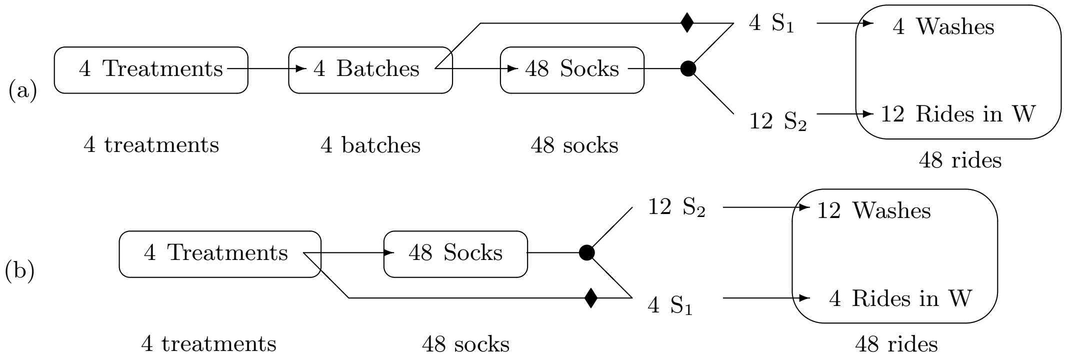

Fig. 22. Example 14: (a) method I; (b) method III

In methods II and III, Treatments are randomized to Socks. Thus socks have individual application of the treatments. In method II, Treatments are again confounded with Washes, with no appropriate residual. In method III, there are twelve washes, each with one sock for each treatment: see Fig. 22(b). Treatments are confounded with Rides within Washes and there is an appropriate residual.

As a compromise between the economy of method I and the estimability of variance in method III, Cox recommends randomly dividing the socks into twelve batches of four, applying each treatment to a whole batch at a time, three batches per treatment, and then "arranging the runs of the simulation machine [Washes] by Method III". What exactly does he mean? If he means Batches are to be handled like Treatments in method III and randomized to Rides within Washes, then this is simple to do, and confounds Treatments with Batches, which itself must be confounded with 12 Rides within four Washes. On the other hand, if he means keep 12 Washes with four rides in each, we need to take care over the potential confounding between Batches and Washes. One way is given in Fig. 23(a), which confounds Treatments with the more variable Washes as well as with Batches. Fig. 23(b) shows another possibility, if batches can be split up for the wear simulation. The pseudofactor B_1 arranges the Batches into four lots, one per treatment. Levels of B_2 are allocated randomly within each level of B_1. The pseudofactor B_2 forms three groups of batches and aliases them with three groups of washes indexed by S_2; each wash has one sock from each batch in the same group, these being allocated to rides in a random order. Now Treatments are confounded with the less variable Rides within Washes, as well as with Batches, although with fewer residual degrees of freedom.

In all cases the randomization of socks is randomized inclusive, with treatments, batches and socks forming a pseudotier.

7. The analysis of multitiered experiments

As described in Brien and Payne (1999), multitiered experiments can be analysed using a generalization of the algorithm implemented in Genstat (Payne et al., 2000). Currently, a more feasible alternative is to use software based on mixed models. In this section we

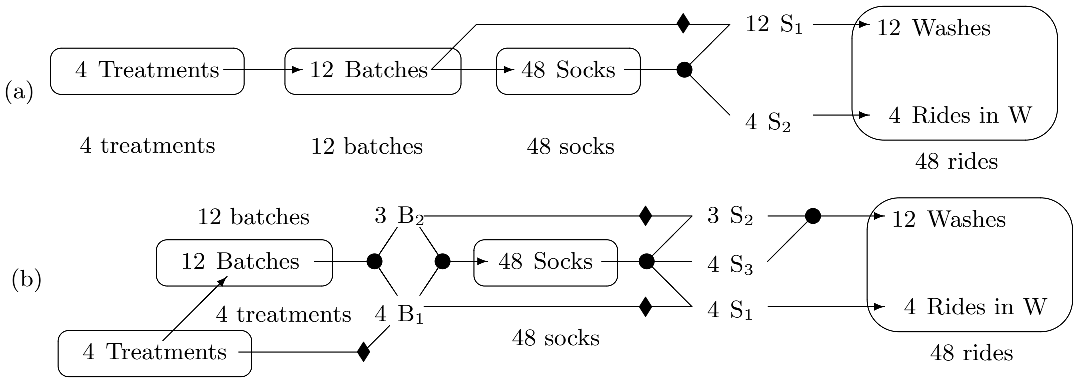

Fig. 23. Example 14: two possibilities for method IV

describe how to use such software for the analysis of multitiered experiments employing, as an example, the two-phase sensory experiment described by Brien and Payne (1999).

Fig. 24 gives the strategy that we propose for formulating the required mixed model. It has similarities to the procedure described by Piepho, Büchse and Emrich (2003).

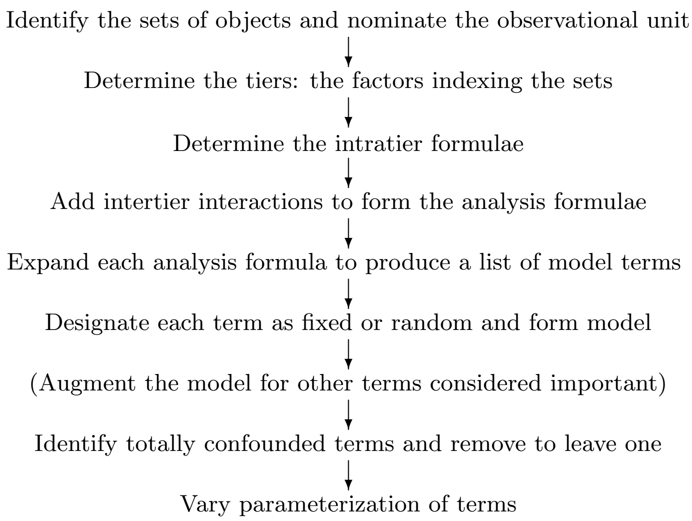

Fig. 24. Steps for obtaining a randomization-based mixed model

The material presented so far in this paper covers the first two steps. The third step is to obtain the *intratier formulae* from the tiers. These formulae represent the crossing and nesting relationships between the factors *within* each tier, given the randomization that has been employed in the experiment. As discussed in Section 2 and in Brien (1983), these relationships will only be the same as the inherent relationships between the factors when the design employed preserves the inherent relationships. We use the notation of Wilkinson and Rogers (1973) to indicate the crossing and nesting relations within each tier. So / and $*$ indicate whether two factors are nested or crossed, respectively. In some instances it may also be necessary to use + to indicate that two factors are additive.

A mixed model in which the terms are derived from the intratier formulae and where all terms are considered random except for those that have never been unrandomized is equivalent to a randomization model (Section 2). However, in experiments in which intertier interactions are anticipated (see Section 7.1), it is necessary to include these interactions by adding the terms to the formula for the highest tier (leftmost in a randomization diagram) in which the any of the interaction's factors occur. The resulting formulae we refer to as the *analysis formulae* to distinguish them from the intratier formulae.

In the next two steps we obtain a model containing fixed and random terms, each term being a generalized factor. The terms can also be obtained manually directly from the randomization diagram. For efficiency we use only a single letter for each factor. We also use Patterson's (1997) convention of listing the fixed terms followed by the random terms; the terms of the same type are separated by plus signs and the two lists by a colon. In addition, as in Piepho, Büchse and Emrich (2003), we underline any unit terms, i.e. terms whose factors uniquely index the observational units.

Next, we consider further augmenting the model to include some or all of the terms that would be justified by the inherent relationships between the factors in a tier, even though they had been ignored in the randomization. Of course, this is generally undesirable, as it results in uncontrolled nonorthogonality that may lead to very poor precision. The preferred practice is to make the design relationships the same as the inherent ones. However, a pragmatic statistician should always be prepared to include, in the analysis, terms for sources of variation that were not anticipated when the experiment was designed.

In addition, because of the nature of the algorithm used in mixed-model software, it is necessary to make sure that no term in the model is inestimable due to total confounding.

Finally, one might want to vary the parametrization of the model terms. For example, the parametrization of fixed terms might be changed from effects to some kind of smooth trend or that of random terms to allow for heterogeneous variances or autocorrelation.

Example 15 (a two-phase sensory experiment): Brien and Payne (1999) describe a two-phase sensory experiment, of which the first, or field, phase is a viticultural experiment and the second, or evaluation, phase involves the assessment of wine made from the produce of the first phase plots. In the field phase, two adjacent Youden squares are used to assign trellis treatments to the plots, a plot being a row-column combination within a square. Each plot is divided into two halfplots and two methods of pruning assigned at random to them. In the second phase, the halfplots from the field phase are randomized, using two Latin squares and an extended Youden design, to glasses in positions on a table for evaluation by judges.

Thus there are three sets of objects: halfplots and treatments from the field phase and evaluations added in the evaluation phase; the observational units are the evaluations. The tiers are given by Brien and Payne (1999) and Fig. 25 shows the randomization diagram based on them. The tiers give the intratier formulae reported in Brien and Payne (1999).

In sensory experiments such as this, it often happens that there is an interaction between Judges and the treatment factors. To allow for this, we modify the last intratier formula by including the factor Judges, but omitting the number of levels for Judges to highlight that it is not from this tier. Thus the analysis formulae are:

$$((2 \text{ Occasions} / 3 \text{ Intervals} / 4 \text{ Sittings}) * 6 \text{ Judges}) / 4 \text{ Positions}$$
$$(3 \text{ Rows} * (2 \text{ Squares} / 4 \text{ Columns})) / 2 \text{ Halfplots}$$
$$4 \text{ Trellis} * 2 \text{ Methods} * \text{Judges}.$$

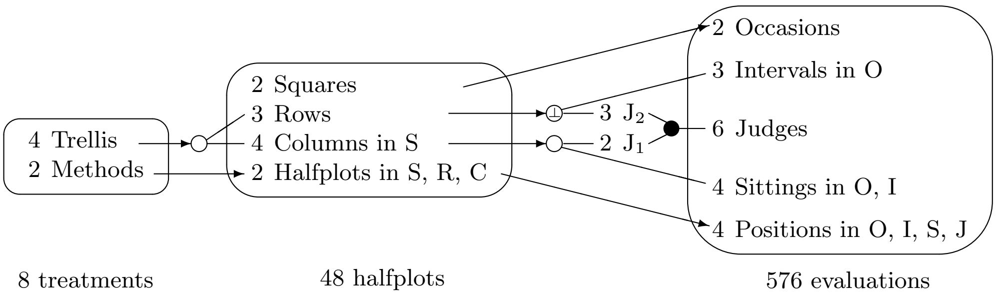

Fig. 25. Randomization diagram for example 15

Next, to decide whether model terms are fixed or random, we designate each factor as fixed or random. The fixed terms are those whose generalized factors comprise only fixed factors; all other terms are random. In this example, it seems reasonable to have the two treatments factors as fixed and all the halfplots and evaluations factors as random with the possible exception of Occasions and Judges. We will take Occasions to be random and Judges to be fixed for the purpose of this paper, recognizing that this is not the only possibility. Consequently our mixed model, with each factor represented by its first letter except for Squares which is represented by Q, is

$$\mathrm{T} + \mathrm{M} + \mathrm{T}\wedge\mathrm{M} + \mathrm{J} + \mathrm{T}\wedge\mathrm{J} + \mathrm{M}\wedge\mathrm{J} + \mathrm{T}\wedge\mathrm{M}\wedge\mathrm{J} : \mathrm{O} + \mathrm{O}\wedge\mathrm{I} + \mathrm{O}\wedge\mathrm{I}\wedge\mathrm{S} + \mathrm{O}\wedge\mathrm{J} + \mathrm{O}\wedge\mathrm{I}\wedge\mathrm{J} + \mathrm{O}\wedge\mathrm{I}\wedge\mathrm{S}\wedge\mathrm{J} + \underline{\mathrm{O}\wedge\mathrm{I}\wedge\mathrm{S}\wedge\mathrm{J}\wedge\mathrm{P}} + \mathrm{Q} + \mathrm{R} + \mathrm{Q}\wedge\mathrm{R} + \mathrm{Q}\wedge\mathrm{C} + \mathrm{Q}\wedge\mathrm{R}\wedge\mathrm{C} + \mathrm{Q}\wedge\mathrm{R}\wedge\mathrm{C}\wedge\mathrm{H}.$$

We are not aware of the need to augment the model for this example. However, the wines from the plots from one square were evaluated on one occasion and those from the second square on a second occasion. As a result the terms O and Q are totally confounded and one must be removed from the model. Hence a final mixed model for initial fitting is:

$$\mathrm{T} + \mathrm{M} + \mathrm{T}\wedge\mathrm{M} + \mathrm{J} + \mathrm{T}\wedge\mathrm{J} + \mathrm{M}\wedge\mathrm{J} + \mathrm{T}\wedge\mathrm{M}\wedge\mathrm{J} : \mathrm{O} + \mathrm{O}\wedge\mathrm{I} + \mathrm{O}\wedge\mathrm{I}\wedge\mathrm{S} + \mathrm{O}\wedge\mathrm{J} + \mathrm{O}\wedge\mathrm{I}\wedge\mathrm{J} + \mathrm{O}\wedge\mathrm{I}\wedge\mathrm{S}\wedge\mathrm{J} + \underline{\mathrm{O}\wedge\mathrm{I}\wedge\mathrm{S}\wedge\mathrm{J}\wedge\mathrm{P}} + \mathrm{R} + \mathrm{Q}\wedge\mathrm{R} + \mathrm{Q}\wedge\mathrm{C} + \mathrm{Q}\wedge\mathrm{R}\wedge\mathrm{C} + \mathrm{Q}\wedge\mathrm{R}\wedge\mathrm{C}\wedge\mathrm{H}.$$

Finally, we might consider a mixed model that allows for unequal variances for the judges.

7.1. Intertier interactions

The term *intertier interaction* refers to an interaction involving factors from more than one tier. It includes block-treatment and unit-treatment interactions (Kempthorne, 1952, Chapter 8; Hinkelmann and Kempthorne, 1994, Section 9.6) and experiment-treatment interactions (Cochran and Cox, 1957, Chapter 14; Cox and Reid, 2000). It is common practice to omit block-treatment and unit-treatment interactions. One reason for this is that often the experiment does not allow their separate estimation. For example, in the randomized complete block design, it is not possible to separate block-treatment, or unit-treatment, interactions from differences between plots within blocks as these sources are totally confounded with each other. A second reason is that, where it is possible to separate block-treatment interactions, a randomization-based analysis does not provide a test for them. A third is that the existence of block-treatment interaction makes inference about overall treatment effects problematic.

However, there are quite a number of experiments where intertier interactions can be estimated and it is highly desirable to do so. We identify two basic types of intertier interaction: randomized-randomized and unrandomized-randomized. Experiments in which treatment factors are randomized using different randomizations will usually involve intertier interactions—the plant experiment in Example 5 is of this kind in that there could be an interaction between Varieties and Regimes which should be allowed for in the model. As both factors involved in the intertier interaction are randomized this is called a *randomized-randomized intertier interaction.* Intertier interactions are also necessary when the experiment involves an unrandomized factor such as Sites, Centres, Laboratories, Sex or Judges (Brien, 1983; Cox and Reid, 2000; Preece, 2001). These factors are likely to interact with the randomized treatment factors and so the corresponding interactions should be included in the model. Such interactions are called *unrandomized-randomized intertier interactions.* The (block-treatment) interactions of Judges with the treatment factors in the two-phase sensory experiment in Example 15 are of this type. Similarly, the two-phase corn-seed experiment in Example 12 may well involve the unrandomized-randomized intertier (experiment-treatment) interaction of Sites with Harvesters.

If intertier interactions are to be included, then they need to be taken into account in designing the experiment. Even though one may employ a design with particular properties in each randomization, this does not necessarily guarantee that the intertier interactions will inherit similar properties. In most of the examples in this paper, where intertier interactions are of interest, the individual designs are orthogonal and the overall structure remains orthogonal when the plausible intertier interactions are included. A counter-example is provided by Example 15, where the Trellis#Judges interaction is not balanced, even though the structure without intertier interactions specified by Brien and Payne (1999) is balanced.

8. Discussion

8.1. Designing Experiments Using Multiple Randomizations

The function of a randomization is to confound terms from the randomized tier with certain terms from the unrandomized tier. Where any arrow goes to a factor which is not nested in all other factors in that tier, as in Figs 3(a), 4 and 5, this should alert the experimenter to the fact that the factor at the tail of the arrow will be estimated with larger variance and fewer residual degrees of freedom than might be possible with a different design. Is such an allocation a physical necessity, or is it a convenience to be paid for in poor precision? Multiple randomizations have the same function, and the different types provide different methods for controlling the confounding.

(a) *Composed* is used where all information from each randomized term is to be confounded with the same term(s) and neither randomization needs information from the outcome of the other; it occurs in two-phase experiments and grazing trials.
(b) *Randomized inclusive* is used where information from a term is to be subdivided according to results of a first randomization and different parts are to be confounded with different terms; it occurs in two-phase and multistage same-unit experiments.
(c) *Unrandomized inclusive* is used where unrandomized factors for second randomization have to take into account the assignment all factors from first randomization; it occurs in superimposed experiments and two-phase experiments.
(d) *Coincident* is used where two terms from different randomizations are to be confounded with the same term(s), and hence with each other, e.g. when treatments in-

troduced in a laboratory phase and replicates from a field phase are to be confounded with same laboratory phase terms; it occurs in a range of experiments including two-phase, single-stage and multistage same-unit experiments.

(e) *Double* is a degenerate type; rotational grazing trials are the only known examples.

(f) *Independent* is used to randomize terms from different tiers to different terms from same tier; it occurs in a range of experiments including two-phase, superimposed and multistage same-unit experiments.

The confounding that occurs as a result of the randomization can be conveniently exhibited in an analysis-of-variance table, as Brien and Payne (1999) do for example 15. However, in order to be able to fully evaluate a design it is necessary to know more about its properties. In Brien and Bailey (2006) we investigate the decomposition of the data space determined by the randomization employed.

8.2. Pseudofactors

Pseudofactors (Yates, 1936; Monod and Bailey, 1992) group the levels of a factor, usually into equally-sized groups, with the feature that there is no scientific rationale for the grouping, but rather it is done as an aid in the design and analysis of the experiment. Their importance in the design of experiments is that they provide a mechanism both for randomizing groups of the levels of a factor, corresponding to the pseudofactor, so that the factor is randomized to more than one factor in another (pseudo)tier and also for identifying groups of the levels of a factor to which the levels of another (pseudo)factor are to be randomized. It seems that this phenomenon occurs more frequently in multitiered experiments than it does in two-tiered experiments and so pseudofactors are especially useful when multiple randomizations are performed.

The grouping of the levels of a factor to form a pseudofactor may be determined in at least three different ways.

(a) The groups are purposefully chosen so that the resulting subspace has certain desirable properties. The L pseudofactors for Lines in Example 9 and the J pseudofactors for Judges in Example 15 are of this type.

(b) The groups reflect the randomization of factors to the factor from which the pseudofactor is derived. Pseudofactors of this type are those for Plots in Example 9 where the Lines, that are randomized to the Plots, determine the pseudofactors. For this type, the randomization diagram has a diamond between the pseudofactor and factors determining it.

(c) The groups which determine the pseudofactors are randomly chosen, deliberately avoiding systematic bias. The pseudofactors for Fibres in Example 4 and for Seedlings in Example 5 are of this type.

Pseudofactors are not incorporated into a tier because, unlike ordinary factors, they are not directly related to real features of the experiment and do not influence the number of objects in a set.

8.3. Factors of intrinsic interest

Another possible dichotomy of the factors is into specific, interesting or predictive factors versus nonspecific, nuisance or standardizing factors (for some discussion of this see Cox, 1958, Section 6.3; Lane and Nelder, 1982; Cox and Reid, 2000; and Preece, 2001). While we agree that this is an important distinction, we contend that is independent of tiers and is relevant for prediction (in the sense of Lane and Nelder, 1982), rather than for model fitting and selection. Examples 5 and 15 demonstrate the independence in that both unrandomized and randomized factors can be specific and nonspecific. In Example 5, Seedlings and Varieties are both randomized factors and in Example 15 both Intervals and Judges are unrandomized. In each case, the first factor is nonspecific and the second is or could be specific.

We believe that the criterion for including the interaction of, say, Judge and Methods in Example 15 should be whether it is anticipated that, in practice, such an interaction could occur. That is, the fact that it is an interaction involving two specific factors plays no role in considering whether to include a term for it in the model. Any potential substantial interactions, whether inter- or intra-tier, need to be in the model.

However, when it comes to prediction, using the selected model, then whether a factor is specific or nonspecific is of central importance and this needs to be determined for each factor, separately from any other classification of the factors. As our prime concern in this paper is with the design of the experiments and model fitting and selection, we have not dealt with specific versus nonspecific factors in any detail. We do observe that predictions will be made for each level of each specific factor but will be averaged over the levels of nonspecific factors.

8.4. Direction of randomization

In Section 2 we suggested that there was a natural direction to a randomization: treatments to plots. While it is, of course, possible to actually achieve the randomization in either direction, it is not possible to randomize plots to treatments by a single permutation of the treatments. Also, we take the view that the basic objective of the randomization is to have the treatments assigned to plots, whichever way that it is accomplished.

In general, the order of the randomization can be established by first identifying the *observational unit*: the smallest unit from which a single value of the response variable is to be obtained. Then, the first set of objects is the set of observational units, Ω, and it is the only set of objects that are never randomized in the experiment. The tier for this set of objects, $\mathcal{F}_\Omega$, consists of the unrandomized factors—the factors whose levels are not associated with particular observational units by randomization.

Once Ω is determined, the next set of objects, Υ, is the one that is randomized first to Ω. After that comes Γ that is randomized either to Υ or Ω. That is, the order of the sets, where possible, follows the ordering inherent in the randomizations.

Example 5 shows that sometimes the direction of randomization is arbitrary. Here one could randomize either seedlings to positions or vice versa. We take the view that the observational units are positions because these physically determine the layout of the experiment and this determines the direction of the randomization.

The direction of randomization is also not well defined when the observational unit differs between the response variables. For example, take grazing trials (Brien and Demétrio, 1998) and, in particular, the continuous grazing experiment in Example 3. Animal response variables, such as weight gain, have a sheep as the observational unit. On the other hand, field response variables, like dry matter production, have a plot as the observational unit. The sets of objects for a field response variable for example 3 are plots, animals and treatments, in this order. It is clear that the observational units are plots and that Blocks and Blocks $\wedge$ Plots are the factors inherent to these units. However, the ordering of the objects for the field response variable has plots and animals reversed compared to that for the animal response variable. It clearly does not follow the randomization order which had treatments randomized to plots and plots randomized to animals. A further difficulty in establishing the order for this example is that it is perfectly feasible to randomize the treatments and animals to plots. This would mean the proposed order of the sets of objects for the field response variable follows the randomization, whereas that for the animal response variable does not. However, when there are more animals than plots it is not possible to do this by applying a single permutation as described in Section 2. Since there must be at least as many animals as plots, to be able to always achieve the randomization by a single permutation requires that plots must be randomized to animals.

It seems that it is necessary to be able to reverse the order of animal and field objects in such trials. The question that arises is: under what circumstances is this possible? The answer seems to be that the original design and the dual of the design employed in randomizing the plots to the animals must both be balanced. This is certainly true for orthogonal designs (pseudofactors may be needed for animals), but is true only for a very limited range of nonorthogonal designs.

8.5. Single Randomizations Viewed as Multiple Randomizations

In some cases one can view the single randomization of a set of factors as multiple randomizations. For example, in a factorial experiment laid out using a randomized complete block design as in Example 2, the randomization of the sets of treatment factors can be regarded as unrandomized inclusive multiple randomizations. The set of treatment factors is divided into two sets, say $\mathcal{F}_\Upsilon$ and $\mathcal{F}_\Gamma$ with associated sets of objects Υ and Γ. Then, Υ, say, is randomized onto Ω after which Γ is randomized to Ω, with tier $\mathcal{F}_\Upsilon \cup \mathcal{F}_\Omega$. Clearly, the division of the treatment factors is arbitrary and somewhat artificial—the randomizations can be achieved in a single randomization.

Another case is that of split-plot experiments, when all treatments have been decided at the outset of the experiment. One might think of the randomization of main-plot treatments as one randomization and that of subplot treatments as another, as in Koch, Elashoff and Amara (1988). However, the independent randomizations can be achieved as a single randomization; i.e. the choice of a single permutation of the units that results in a random association between the complete set of subplot units and all the treatment factors.

Whenever apparently multiple randomizations can be achieved by selecting a single permutation, they will be regarded as a single randomization. Such a randomization can always be adequately described using two tiers.

References

Bailey, R. A. (1981). A unified approach to design of experiments. *J. R. Statist. Soc.* A, *144*, 214–223.

Bailey, R. A. (1989). Designs: mappings between structured sets. In J. Siemons (Ed.), *Surveys in Combinatorics, 1989*, Volume 141 of *London Mathematical Society Lecture Notes Series*, pp. 22–51. Cambridge: Cambridge University Press.

Bailey, R. A. (1991). Strata for randomized experiments. *J. R. Statist. Soc.* B, *53*, 27–78.

Bailey, R. A. (1996). Orthogonal partitions in designed experiments. *Des. Codes Cryptogr., 8*, 45–77.

Box, G. E. P. and S. Jones (1992). Split-plot designs for robust product experimentation. *J. Appl. Statist., 19*, 3–26.
Brien, C. J. (1983). Analysis of variance tables based on experimental structure. *Biometrics, 39*, 51–59.
Brien, C. J. (1992). *Factorial Linear Model Analysis.* Ph. D. thesis, Department of Plant Science, The University of Adelaide, Adelaide, South Australia. URL: <http://chris.brien.name/thesis>.
Brien, C. J. and R. A. Bailey (2006). Decomposition tables for multitiered experiments. submitted for publication.
Brien, C. J. and C. G. B. Demétrio (1998). Using the randomisation in specifying the ANOVA model and table for properly and improperly replicated grazing trials. *Aust. J. Exp. Agric., 38*, 325–334.
Brien, C. J., P. May, and O. Mayo (1987). Analysis of judge performance in wine-quality evaluations. *J. Food Sci., 52*, 1273–1279.
Brien, C. J. and R. W. Payne (1999). Tiers, structure formulae and the analysis of complicated experiments. *The Statistician, 48*, 41–52.
Clarke, G. M. and R. E. Kempson (1997). *Introduction to the Design and Analysis of Experiments.* London: Arnold.
Cochran, W. G. and G. M. Cox (1957). *Experimental Designs* (2nd ed.). New York: Wiley.
Cox, D. R. (1958). *Planning of Experiments.* New York: Wiley.
Cox, D. R. and N. Reid (2000). *The Theory of the Design of Experiments.* Boca Raton: Chapman and Hall/CRC.
Cullis, B. R., A. B. Smith, J. F. Panozzo, and P. Lim (2003). Barley malting quality: are we selecting the best? *Aust. J. Agric. Res., 54*, 1261–1275.
Federer, W. T. (1975). The misunderstood split plot. In R. P. Gupta (Ed.), *Applied Statistics*, pp. 9–39. Amsterdam: North Holland.
Fisher, R. A. (1935). Contribution to the discussion of "Complex experiments" by F. Yates. *J. R. Statist. Soc., Supp., 2*, 229–231. URL: <http://www.library.adelaide.edu.au/digitised/fisher/129.pdf>.
Freeman, G. H. (1959). The use of the same experimental material for more than one set of treatments. *Appl. Statist., 8*, 13–20.
Heiberger, R. M. (1989). *Computation for the Analysis of Designed Experiments.* New York: Wiley.
Hinkelmann, K. and O. Kempthorne (1994). *Design and Analysis of Experiments: Volume I: Introduction to Experimental Design.* New York: Wiley.
Kempthorne, O. (1952). *The Design and Analysis of Experiments.* New York: Wiley.
Kempthorne, O. (1955). The randomization theory of experimental inference. *J. Am. Statist. Ass., 50*, 940–967.
Kerr, M. K. (2003). Design considerations for efficient and effective microarray studies. *Biometrics, 59*, 822–828.
Koch, G. G., J. D. Elashoff, and I. A. Amara (1988). Repeated measurements – design and analysis. In S. Kotz and N. L. Johnson (Eds.), *Encyclopaedia of Statistical Sciences*, Volume 8, pp. 46–73. New York: Wiley.
Lane, P. W. and J. A. Nelder (1982). Analysis of covariance and standardization as instances of prediction. *Biometrics, 38*, 613–621.
McIntyre, G. A. (1955). Design and analysis of two phase experiments. *Biometrics, 11*, 324–334.
Mead, R. (1988). *The Design of Experiments: Statistical Principles for Practical Applications.* Cambridge: Cambridge University Press.
Mead, R. and R. N. Curnow (1983). *Statistical Methods in Agriculture and Experimental Biology.* London: Chapman and Hall.
Mee, R. W. and R. L. Bates (1998). Split-lot designs: experiments for multi-stage batch processes. *Technometrics, 40*, 127–140.
Miller, A. (1997). Strip-plot configurations of fractional factorials. *Technometrics, 39*, 153–161.

Monod, H. and R. A. Bailey (1992). Pseudofactors: Normal use to improve design and facilitate analysis. *Appl. Statist., 41*, 317–336.

Nelder, J. A. (1965a). The analysis of randomized experiments with orthogonal block structure. I. Block structure and the null analysis of variance. *Proc. Roy. Soc. A, 283*, 147–162.

Nelder, J. A. (1965b). The analysis of randomized experiments with orthogonal block structure. II. Treatment structure and the general analysis of variance. *Proc. Roy. Soc. A, 283*, 163–178.

Patterson, H. D. (1997). Analyses of series of variety trials. In R. A. Kempton and P. N. Fox (Eds.), *Statistical Methods for Plant Variety Evaluation*, pp. 139–161. London: Chapman & Hall.

Payne, R. W., D. B. Baird, A. R. Gilmour, S. A. Harding, P. W. Lane, D. A. Murray, D. M. Soutar, R. Thompson, A. D. Todd, G. Tunnicliffe Wilson, R. Webster, and S. J. Welham (2000). *The Guide to Genstat, Part 2 Statistics.* Oxford: VSN International Ltd.

Piepho, H. P., A. Büchse, and K. Emrich (2003). A hitchhiker's guide to mixed models for randomized experiments. *J. Agron. & Crop Sci., 189*, 310–322.

Preece, D. A. (1991). Contribution to the discussion of "Strata for randomized experiments" by R. A. Bailey. *J. R. Statist. Soc.* B, *53*, 67–68.

Preece, D. A. (2001). Types of factor in experiments. *J. Statist. Planning & Inf., 95*, 269–282.

Schoen, E. D. (1997). Cheese-making with Genstat: a case study in design of industrial experiments. *Genstat Newsletter, 33*, 20–29.

't Mannetje, L., R. J. Jones, and T. H. Stobbs (1976). Pasture evaluation by grazing experiments. In N. H. Shaw and W. W. Bryan (Eds.), *Tropical Pasture Research: Principles and Methods*, Bulletin 51, pp. 194–234. Slough: Commonwealth Agricultural Bureaux.

Tjur, T. (1984). Analysis of variance models in orthogonal designs. *Internat. Statist. Rev., 52*, 33–81.

White, R. F. (1975). Randomization and the analysis of variance. *Biometrics, 31*, 552–572.

Wilk, M. B. and O. Kempthorne (1957). Non-additivities in the Latin square design. *J. Am. Statist. Ass., 52*, 218–236.

Wilkinson, G. N. and C. E. Rogers (1973). Symbolic description of factorial models for analysis of variance. *Appl. Statist., 22*, 392–399.

Wood, J. T., E. R. Williams, and T. P. Speed (1988). Non-orthogonal block structure in two-phase designs. *Aust. J. Statist., 30A*, 225–237.

Yates, F. (1936). A new method of arranging variety trials involving a large number of varieties. *J. Agric. Sci., 26*, 424–455.

Yates, F. (1975). The early history of experimental design. In J. N. Srivastava (Ed.), *A Survey of Statistical Design and Linear Models*, pp. 581–592. Amsterdam: North Holland.